U.PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

TRACING SEPHARDIC JEWRY THROUGH GENETICS: CRYPTO-JEWS AND THE SECOND DIASPORA

Maria Inês Pires Nogueiro

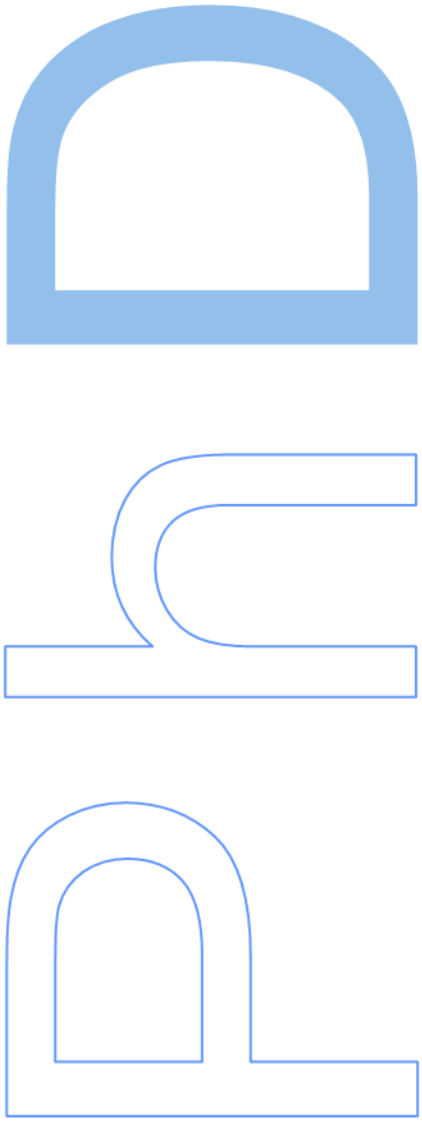
Tese de Doutoramento apresentada à Faculdade de Ciências da Universidade do Porto
Biologia
2015

TRACING SEPHARDIC JEWRY THROUGH GENETICS: CRYPTO-JEWS AND THE SECOND DIASPORA

Maria Inês Pires Nogueiro
Programa Doutoral em Biologia
Biologia
2015

Orientador
Doutora Leonor Gusmão,
Professora Adjunta da Universidade do Estado do Rio de Janeiro (UERJ)
Investigadora Afiliada do Ipatimup

Coorientador
Professor Doutor António Amorim dos Santos,
Professor Catedrático, Faculdade de Ciências da Universidade do Porto



“Your purpose...should always be to know... the whole that was intended to be known.”

Maimonides, *in The Guide for the Perplexed.*

Illumination dated from 1347-1348 of the Hebrew translation Maimonides teaches about the “measure of men”.
Royal Library in Copenhagen.

À minha Gente,

Agradecimentos

No fim destes quatro anos de tanto trabalho, de tanta viagem, de tanta gente, de tanta generosidade, quase parece incrível que tudo se resuma a menos de 200 páginas. Mas são duzentas páginas que levam nelas muitas outras páginas, de muitas outras histórias, e quero agradecer a todos aqueles que, de uma forma ou de outra, me ajudaram a chegar aqui.

Aos meus *Mestres,* Leonor Gusmão e António Amorim:

Leonor, agradeço ter-me cruzado contigo nesta viagem. Obrigada por tudo o que me ensinaste, pelo pragmatismo, pela força e persistência, pela generosidade, pela lucidez, por toda a cumplicidade, mas, sobretudo, pelo exemplo! Quando for "grande", gostava definitivamente de ser uma pessoa como tu!

Professor Amorim, obrigado por me ter recebido no seu grupo, por desde início me ter apoiado neste projeto, por todo o conhecimento, tempo e paciência interminável dispensados, por todos os esclarecimentos, por me ajudar a crescer em todo este processo e ter estado sempre presente. Espero ter estado à altura!

Ao IPATIMUP, e ao Professor Sobrinho Simões, na qualidade de Presidente do instituto, por me ter dado a oportunidade e todas as condições necessárias para desenvolver o meu trabalho de investigação.

À Fundação para a Ciência e Tecnologia (FCT) pelo projeto e bolsa de estudo sem os quais não poderia ter realizado este trabalho.

Ao Luís, muito obrigada por tudo! Pelas horas imensas de cálculos e tabelas e mais gráficos! Por toda a tua generosidade, disponibilidade e amizade. Por me teres "aturado" em todos os momentos de trabalho e conversas mais ou menos filosóficas! *Muchas gracias*!

À Cristina, por toda a cumplicidade e toda a amizade. Por todos os momentos em que estiveste sempre lá! O percurso foi mais fácil contigo por perto! Obrigada pelas gargalhadas e por me aturares nos momentos mais complicados!

À Sofia, por toda a ajuda, sobretudo nesta última fase! Obrigada pela tua generosidade, e todo o empenho que sempre mostraste nas coisas que fazes. Espero sinceramente que consigas continuar a trabalhar no que gostas, porque bem o mereces!

À professora Maria João e todos os membros do grupo de genética populacional, pelo excelente ambiente de trabalho que me proporcionaram sempre. Em especial, à Cíntia, Verónica e Rui por me terem aturado com tantas dúvidas e questões.

Às comunidades Judaicas de Lisboa, Porto, Rio de Janeiro, Recife e Natal por me terem recebido e apoiado de forma tão franca e terem contribuido com as amostras fundamentais para a realização deste trabalho. Em especial, ao João Guerra, rabino Eliezer Shai Di Martino, Ferrão Filipe, Jorge Neves, David Albagli Gorodicht, Isaac Kayat, Tânia Neumann Kaufman, Nivaldo Vasconcelos, Rabino Yohanan Yedidya Medeiros, Ezequiel Pessaj, Cleudo Freire, Brenno Brummel. À Nicole Cohen-addad pelas amostras de São Tomé.

A todos os que colaboraram com amostras em Bragança, Carção, Argozelo, Mogadouro e Vilarinho dos Galegos, em particular ao Paulo Lopes, Carlos Prada, Francisco Pires, Filipe Campos.

Ao Jorge Martins, Professora Elvira Mea e Fernanda Guimarães por toda a ajuda nas questões históricas.

Ao Kevin, pela amizade e tempo dispensados em correções. O meu “inglês” ficou definitivamente melhor!

Aos meus amigos de sempre, em especial ao João e Ana, por estarem sempre lá.

Ao Eduardo, por me ter acompanhado sempre, em todos os momentos.

À minha família, por serem o meu alicerce e a minha força.

Um grande bem haja a todos!

Table of Contents

Resumo

Cripto-judaísmo, *stricto sensu*, é definido como a adesão secreta ao judaísmo enquanto publicamente é professada outra fé. Sendo o resultado geral da intolerância religiosa, tem sido mais especificamente associado com a Inquisição Ibérica, tanto na Europa como nas colónias portuguesas e espanholas ultramarinas. A persistência desta identidade, por mais de quatro séculos, é, em si, um tema fascinante de pesquisa histórica e antropológica. Esclarecer em que medida este fenómeno cultural poderia estar refletido no património genético dos descendentes de Judeus, foi o principal objetivo do presente trabalho.

Para o efeito, foram analisados marcadores genéticos capazes de distinguir as histórias das linhagens paternas e maternas (cromossoma Y e mtDNA, respectivamente), assim como marcadores recombinantes, autossómicos e ligados ao cromossoma X. As comunidades que permaneceram na Península Ibérica, designadamente as do NE de Portugal, foram utilizadas como representantes da população ancestral, comparativamente aquelas que, após o Decreto de Expulsão, migraram para o Novo Mundo (Brasil) e África (São Tomé).

Relativamente aos marcadores uniparentais, tanto as estimativas de níveis de diversidade como do número de efectivo populacional feminino, apontaram para um tamanho estável na população Judaica portuguesa, de acordo com os resultados anteriores obtidos para a componente masculina. Traços Sefarditas, bem como sinais de introgressão a partir da população não-judaica envolvente, foram também identificados. Em conformidade com os resultados obtidos para marcadores uniparentais, foram detectados níveis altos de diversidade genética no genoma recombinante. A resistência ao declínio esperado na diversidade, resultante de deriva genética, foi provavelmente compensada pela miscigenação com populações não Judias. A análise simultânea de marcadores autossómicos e heterossómicos, revelou uma tendência de género na história demográfica desta população, sugerindo uma assimetria entre os tamanhos efetivos da população feminina e masculina no processo de miscigenação entre judeus e não-judeus.

Relativamente aos presumíveis descendentes de Judeus de São Tomé, a diversidade genética encontrada foi semelhante à anteriormente reportada para a população geral não Judaica. Foi detectada uma forte componente Africana, mas sem sinais de ancestralidade putativamente Judaicos.

No que respeita às comunidades Judaicas Brasileiras, foram encontrados niveis de diversidade genética altos e um padrão demográfico assimétrico evidenciando um excesso desproporcional

de migrantes do sexo masculino, juntamente a uma forte introgressão da população nativa e população africana feminina na referida comunidade.

Palavras-chave: cripto-judeus; marcadores uniparentais; genoma recombinante; comunidades da Diáspora.

Summary

Crypto-Judaism, stricto senso, is defined as the secret adherence to Judaism while publicly professing another faith. A general outcome of religious intolerance, it has been more specifically associated with Iberian Inquisition, both in Europe and in Portuguese and Spanish colonies overseas. The persistence of this identity, for over four centuries, is in itself a fascinating historical and anthropological research topic. To clarify to what extent this cultural phenomenon could be reflected in the Jewish descendants `genetic portrait was the main focus of the present work.

For this purpose, genetic markers able to disentangle the paternal and maternal lineages histories (Y chromosome and mitochondrial DNA, respectively) as well as autosomal and X-linked non-coding markers were analyzed.

The communities that stayed in Iberia, specifically those from the NE Portugal were used as proxy to the ancestral population, to those that went to exile to the New World (Brazil) and Africa (São Tomé) after the Expulsion Edict.

Regarding uniparental markers, estimates of both the diversity levels and the number of female effective-population founders for the Portuguese Jews, pointed at a stable size of the studied population, in agreement with previous findings for the male counterparts. Sephardic signatures as well as signs of introgression from the host non-Jewish population were also identified. Confirming the results from lineage markers, high levels of genetic diversity at the recombining genome were detected. The resistance to expected diversity decline due to genetic drift was probably compensated by admixture with the non-Jews. The simultaneous analysis of both autosomal and heterosomal markers revealed a sex biased demographic history, suggesting an asymmetry between female and male effective population sizes in the admixture process between Jews and non-Jews.

Concerning the supposed Jewish descendants from São Tomé, the genetic diversity was at the same level of that previously reported for the general non-Jewish population. No signs of the putative Jewish lineages were found, instead, high percentage of African ancestry was detected.

The Diaspora Jews in Brazil showed high levels of genetic diversities and an asymmetric, strongly sex-biased pattern: a disproportionate excess of male migrants, along with a strong introgression of Native and African females.

Key words: Crypto-Jews; uniparental markers; recombining genome; Diaspora communities

List of Abbreviations

AIMs Ancestry Informative Markers

BCE Before Common Era

bp base pair

CE Common Era

dbSNP SNP database

D-loop Displacement loop

DNA DeoxyriboNucleic Acid

Indels Insertions/deletions

Kb kilobase

LD Linkage Disequilibrium

MSY Male Specific region of Y chromosome

mtDNA mitochondrial DNA

NCBI National Center for Biotechnology Information

Ne effective population size

NRPY Non-Recombining Portion of Y chromosome

NRY Non-Recombining region of Y chromosome

PARs Pseudo Autosomal Regions (PAR1/PAR2)

PCR Polymerase Chain Reaction

PMRCA Place of the Most Recent Common Ancestor

rCRS revised Cambridge Reference Sequence

RFLP Restriction Fragment Length Polymorphism

RNA RiboNucleic Acid

RSRS Reconstructed Sapiens Reference Sequence

SNP Single Nucleotide Polymorphisms

SRY Sex-determining Region of Y chromosome

SSRs Short Sequence Repeats

STR Short Tandem Repeat

TMRCA Time to Most Recent Common Ancestor

tRNA transfer RNA

Figures and Tables Index

Chapter I. Introduction

Chapter IV: Preliminary results of the Diaspora Jews in Brazil

Thesis foreword

The main goal of this work was to contribute, with a population genetics perspective, to the historical and anthropological knowledge concerning the Sephardic Jewish population.

It is well known that a cultural cleavage between different fields of knowledge, particularly between humanities and natural sciences, has existed for longer than is desirable. A multidisciplinary methodology for the research of a specific topic supposes an integrative dialogue and an understanding of numerous standpoints, neither of which are always easy to accomplish. Nevertheless, in the present thesis, such an approach was undertaken.

In the following, an introduction to the historical context of the Jewish population as well as to the basic principles of population genetics is addressed. The purpose of this contextualization is to call upon both sciences, and by doing so, give the necessary background from both fields of knowledge for a broader understanding of Jewish population history.

The first part of this thesis (Introduction) begins with a historical summary of the Jewish populations, where the complexity of their demographic past and the nuances of identity which are owed to their troubled history are shown.

The second part of the introduction is devoted to Population Genetics as a research tool for the survey and construction of population complex demographic histories. Special attention is given to the processes that shape the genetic pool, particularly those that act in the modulation of genetic diversity patterns. Recombining (autosomes and X chromosome) and non-recombining (Y chromosome and mitochondrial DNA) portions of the human genome are referred to in the context of molecular markers, that allow access to the genetic portrayal of a population.

The results are presented as scientific papers. Population genetic data are reported for different genetic marker sets in order to obtain a comprehensive history of the Sephardic Jews. Preliminary results for the Diaspora Jewish community of Brazil are also presented.

Finally, in the discussion section, an integrative balance of the most relevant results and conclusions is addressed, as well as future research directions. This includes the high diversity indices found, signs of sex-biased admixture with respective host populations (European for Portuguese Jews and Native American and African for the diaspora populations), and also the presence of an ancestral putative Jewish background.

CHAPTER I. INTRODUCTION

A Population: The Jews

Defining Jewishness

Defining who is a Jew is not a straightforward task. According to the on-line version of the Oxford dictionary, a Jew is described as "*a member of the people and cultural community whose traditional religion is Judaism and who trace their origins to the ancient Hebrew people of Israel*". What at a first glance could seem a simple definition includes in fact myriad of complex concepts of religion, history, ethnicity, culture, genealogy and identity. This definition assumes particular importance when a study is focused on Iberian Jews and the Iberian Diaspora communities. Several terms have been used in historical, anthropological and genetic works to designate the Iberian Jews or their descendants, such as *Conversos*, *New-Christians*, *Anusim, Crypto-Jews or Marranos*. Often these words are used as synonymous, since they all characterize those who were forcedly converted to Christianity. However, they can also represent dissimilar identities and attitudes towards Judaism. In a first category, the *Conversos* and *New-Christians* usually denote those who eventually were integrated in the general Christian population [1-3]. As a second category, the *Anusim, Crypto-Jews* and *Marranos* could be grouped, and these are the terms used for those who kept *Judaizing*, which means following the Mosaic Law (from Moses) in secret. Though, the two categories often present a rather ill-defined boundary between them. In fact, this phenomenon is more complex than the simple hiding of a proscribed and persecuted identity, involving the persistence of a more popular or possible Judaism instead of a more formal or Rabbinical religious attendance for long periods of time in a fairly heterogeneous historical and geographical context. It contains elements of religious syncretism in its construction and at the same time, a tendency to maintain a distance and isolation from the surrounding Christian culture. Ultimately, this could be considered as a means of maintaining a group identity through the remembrance of a common origin rather than a theological or a strict religious practice in its essence [1-3].

The subject Jewish communities of the present study are included in the latter category, Crypto-Jews or Marranos; thus their Jewishness is impregnated with paradox and a duality of identity between Judaism and Christianity. Their struggle for survival as individuals and also as a collective identity for centuries bears the brands of its troubled history.

History of the Jewish People

Origin

To understand the collective memory of a people it is essential to know their history. Again the words Hebrews, Israelis or Jews are usually indifferently used, but probably the most embracing term would be “Jews” (from Judah) although all others also have a justification. Like any ancient people, tracing the Jews´ origin and evolution is rather complex. Frequently it is difficult to define the diffuse boundary between the real facts from those who belong to the field of uncertainty and myth [4]. The story of the ancient Hebrews is based on the Pentateuch, an anthology that includes five books which the Jews call the Torah, meaning the Law [5].

According to the first book, the Genesis, the Hebrews are a Semitic people descended from Abraham of Canaan. Abraham is the first of the “*Three Patriarchs*”. His son Isaac, the second patriarch, had two sons, Esau and Jacob. Jacob is the last patriarch and his twelve sons gave rise to the twelve tribes of Israel: Judah, Levi, Asher, Benjamin, Dan, Gad, Issachar, Joseph, Naphtali, Reuben, Zebulon and Simeon [4, 5]. Jacob was also called Israel, and his sons, “sons of Israel” or Israelis [5]. This is known as the *Patriarchal Age*, which probably occurred between the 20th to the 16th centuries BCE [4].

Archaeological evidence actually points to the emergence of the Kingdom of Israel in the early Bronze Age, associated with the Canaanite culture, supported by technological, linguistic and ethnological characteristics [6]. Information from the Nuzi texts, Mari and Alalakh documents, Egyptian Execration tablets and the Amarna letters have been used to support the idea of the Patriarchs as historical figures who migrated with semi-nomadic people from Mesopotamia to Canaan [7]. Before the *Patriarchal Age*, the Israelis were known as Hebrews, according to the Jewish encyclopedia, from the name “Eber” an ancestor of Abraham, but the name could also refer to those from “the province beyond" the Euphrates. A more recent and controversial justification links the Hebrews to the Habiru people, a semi-nomadic group described as rebels, outlaws, bowmen, servants, slaves or migrant laborers [7].

Some of the Israelite tribes settled in Egypt between the 16th - 13th centuries BCE. The period of the monarchy begins with Saul (1029-1007 BCE) from the tribe of Benjamin, but it was the king David (1007-970 BCE) who unified the twelve tribes and conquered Jerusalem. For the following 1200 years, the land of Israel witnessed the successive territorial domain of the Assyrians, Persians, Greco-Macedonian and the Roman Empire [4]. From the initial ethnic unity very little survived from these twelve centuries of constant deportations, migrations, conquests,

annexations and territorial losses [8]. This tribal disintegration did not affect, however, the religion, the tradition, the language and the reference to a Jewish origin [8]. Thus, the current Jews can be divided into four major groups depending on the geographical region of the globe they settled in: the Ashkenazim, Jews from France, Germany and Eastern Europe; Sephardim, from the Iberian Peninsula and North Africa; Mizrakhim or Oriental Jews from the Middle East and the Falasha group from Ethiopia [6].

Iberian/Portuguese Jews

Jewish diaspora towards Iberia certainly occurred a long time ago, although the exact date of their arrival is still today unsettled. According to the Torah, their genealogy traces back to the people who fled from Judea after Nabucodonosor's conquest of Jerusalem, in 587 BCE [9-11]. Christian and Jewish humanists from the Renaissance were both interested in unraveling their origins: some argued that Jews first arrived in the Iberian Peninsula along with the diaspora of The Ten Lost Tribes of Samaria, in 722 BCE [11, 12] while others claimed that the Israeli arrived with the Tarshish's ships constructed by King Solomon around 900 BCE [11] or with the Phoenicians, in 1200 BCE, although this date is not archeologically documented [11, 13].

The oldest archaeological evidence of Jewish presence in Iberia known so far was recently found in the south of Portugal (Silves) and has a chronology of 390 CE (http://www.uni-jena.de/en/News/PM120525_Schrifttafel.html).

Written documents mentioning Jews in Iberia accumulate from the beginning of the Visigoth period onwards: in the 4th century CE, the decisions from the Council of Elvira, particularly the interdiction of marriages between Jews and Christians, confirm their complete integration among Iberian communities, except for the religious cult [11, 12]. These are in fact the first documents relating restrictive measures towards the Jews, and centuries later, they would be reproduced in the canonic laws that structured the Inquisition [14].

The relative relief from ecclesiastical pressure brought by the Visigoths' invasion was brief: first with the catholic conversion of King Recaredo in 587 CE, followed by the decree of King Sisebuto in 613 CE, imposing the expulsion of all Iberian Jews who would not embrace the Christian faith [11, 12]. This episode marks the beginning of Crypto-Judaism and was a harbinger of their subsequent dramatic history in the Iberian Peninsula [11, 12].

The Islamic invasion in 711 caused a division between supporters and opponents of the Muslims among the Judaic community. Indeed, if it seems certain that no Jewish-Muslim alliance occurred before the invasion, the same did not happen after Tarik's troops crossed Gibraltar Strait, and the Jews, once pariahs, had then the opportunity to administer several important regions of the Iberian Peninsula [11].

Over the 11th and 12th centuries, a strong influence of Islamic culture was evident in the Iberian Peninsula, particularly in philosophy, geography, astronomy, mathematics and medicine, providing a cultural blossoming often qualified as the golden age of the Sephardic Jews [14].

From the 12th to the end of the 13th century, the geography of Judaism changed significantly with the Christian settlement policies, to which the first Jewish medieval colonies owe their existence [11]. From the beginning of the Portuguese nation in 1143, till the Expulsion Edit in 1496, the successive Portuguese monarchs would balance the anti-Judaic clerical and popular pressures and the predominance of the Jewish social and economic life [12]. Still, there was a climate of relative tolerance in the coexistence between Jews and Christians, favoring the emergence of numerous Jewish communes [14, 15].

If there were favorable measures towards the Jewish population such as free religious practice, the construction of synagogues, the exemption from certain taxes, the permission to occupy public roles as well as live among the Christians or in the *Judiaria* (Jewish quarters), several opposite restrictive policies were also adopted. Some of these measures included the duty to wear distinctive vesture marks, the imposition to live separately from the Christian neighborhoods, the restriction in access to public posts or the payment of various taxes [9, 12, 14, 15].

The 14th century marked a dramatic change in this fragile equilibrium, which nevertheless had allowed the emergence of Jewish communes in the country. The estimates of the size of the Jewish population vary between 30,000 to more than 60,000. This number would soon increase substantially with the arrival of around 100,000 Spanish Jews fleeing from their country [12, 13, 16].

During the second half of the 14th century, anti-Judaic intolerance was steadily increasing in the neighboring country (Spain), which culminated in the massacres of 1391. In 1478 the Inquisition was established, followed by the Edit of Expulsion by the Catholic Kings. These events would have a strong impact on the future of the Portuguese Jews [12, 14, 17]. The initial tolerance

towards the Spanish exiles and the Portuguese Jews was, however, doomed. The marriage of the Portuguese King Manuel I to the daughter of the Catholic Kings had severe political implications [12, 14, 17]. Thus, in December of 1496 the King signed the Portuguese Edict of Expulsion, ordering the departure of Moors and Jews by October of the following year [14, 17]. However, in May of 1497, about 20,000 Jews from all over the country, preparing to exile, were forcedly baptized. As a result, officially, there were no more Jews in Portugal, instead, a new identity was created: the new-Christians or *Conversos* [12]. This ambiguous policy highlights the socioeconomic importance of the Jews. King Manuel I prohibited inquiries into the Jewish faith for 20 years, consenting to an accepted crypto-Judaism [12, 14].

The Papal Bull establishing the Inquisition in Portugal was issued in 1536 under the reign of King João III. In the 17th and18th centuries, the inquisitorial processes intensified and as a result, there was a significant exodus to other countries, particularly of manufacturers and the merchant elite [9]. A total of 24,522 Jews (including Penitents, Relaxed in Flesh and Relaxed in Effigy) would suffer the unspeakable terror of this Court [12].

By the end of the 15th century there were about 134 Jewish quarters, communes or communities throughout the country, with an estimated population of 100,000 people, which translates into 10% of all the Portuguese population at that time. The exact number of people who emigrated is not known, however it is thought that by 1631 the Jewish population was reduced to approximately 10,000 [16].

The rebirth of the Jewish communities in Portugal took place in early 19th century, when the Marquis de Pombal ended the official discrimination and persecution performed by the Inquisition [12, 16]. The Israeli community of Lisbon was founded by Sephardic Jews from North Africa [12, 14], and in the 20th century, the communities of Porto, Bragança, Belmonte, Faro and the Azores emerged [9, 12-14, 17].

In particular, the communities of Bragança and Belmonte arose thanks to the work of Samuel Schwarz and Captain Barros Basto, with a movement in the early 20th century which aimed to bring back the crypto-Jews to normative Judaism. While in Belmonte the community is still dynamic today, the Bragança community was dissolved in 1934, shortly after its appearance, and its population dispersed in the region. A strong sense of belonging among those Jewish descendants is still well alive today and the complexity of their Marrano identity is well pictured in a medallion from Carção (NE Portugal) dated probably from the 17th century, (**Figure 1**)

where in one face a menorah, the most widespread Judaism symbol is represented and on the other face, the *Agnus Dei*, symbol of Christianity.

.

Figure 1. Double face medallion with a Menorah and *Agnus Dei, the Lamb of God*

Adapted from Tavares (2015) [18]

The Iberian Diaspora

Jewish people are a paradigm of constant geographical mobility over an enormous time span throughout human history and so the stereotype of the “wandering Jew” or the “wandering people” has indeed some substantiation [4, 19]. As a minority, Jewish populations were exposed to legal, socioeconomic, political and cultural constraints and while these influences were present across Jewish communities worldwide, it is also true that significant local differences created unique motivations or constraints to stay or move, often translated in expulsions or deportations along the Jewish history [19].

The major diaspora of Marrano and Crypto-Jews actually begins in the 15th and extends until the 18th century when the legislation of Marquis of Pombal put an end to the official

discrimination and persecution of the Inquisition [12, 16]. Initially, they settled in Amsterdam, London, Hamburg, Turkey, some French and Italian cities, and north Africa; from the mid-sixteenth century, they migrated to the Portuguese colonies in Africa, India and Brazil, and later from the cities of northern Europe to the New World, Curacao, Paramaribo, and the USA, where Jewish colonies founded synagogues with the Portuguese rite [12, 14, 16].

Jewish Diaspora in São Tomé e Príncipe

São Tomé e Príncipe are the two main islands of a small archipelago located in the Gulf of Guinea, of the western equatorial coast of Africa. These islands were probably uninhabited at the time of the Portuguese discovery in 1471, but soon after that the first settlements took place, in 1486 by João de Paiva. It is documented that he took with him many settlers, the great majority of them Jews. One of the most controversial episodes of this colonization is the forced migration of Jewish children in 1492 [20]. The information about this occurrence is scarce and diverges between Jewish and Christian historians from that time, although the reason seems to be unanimous: clandestine Jewish refugees from the Spanish expulsion or those that were allowed to stay in Portugal for a period of 8 months and could not manage to leave the country before the stipulated dates and therefore were forced into slavery. Around 2000 young children and teenagers were taken from their families, forcedly baptized [21] and sent to the archipelago so they could be away from the Jewish faith, grow up with a Christian education and contribute to the population of the new territory, one of the main aims of the Portuguese king Joao III [20]. Several documents confirm that many of these children survived and along with other settlers, mainly of African origin, started the process of miscegenation [22]. After the decree of expulsion in Portugal, many new-Christians also fled to São Tomé e Príncipe, since the Inquisition was never established there [12, 20].

Over more than a century several documents attest to the continuous movement of new-Christians to São Tomé, which worked as a refuge from the inquisitional prosecutions. With the flourish of the intercontinental commercial trade by New Christians, São Tomé became an import path for those who fled from Iberia to South America [1]. Despite the documents above mentioned, the information available on the impact or number of Jewish people that really went (and when) to São Tomé is scarce.

Jewish Diaspora in Brazil

The Portuguese colonization of Brazil was essentially agricultural. Many New-Christians that migrated or fled from the Inquisition to Brazil also became landowners and probably due to their higher level of skills, compared to other migrants, and regardless of several restrictive measures, some of them occupied political and military posts. In 1502, just two years after the arrival of Pedro Alvares Cabral, the first contract of exploration of the Brazilian redwood was attributed to a New-Christian [16]. Brazilian redwood was the first commercial product brought to the European markets through the Jewish merchant network spread throughout Spain, France and the Netherlands. In the mid-16th century, a well establish Jewish transatlantic diaspora begin to emerge as a result of the strong trade connections. At this time sugar cane started to be cultivated in São Tomé, and from Africa, shortly after, also in Brazil. This attracted many Jewish migrants to the New World that concentrated their activities in the international trade not only of sugar and wood but also due to the great expansion of the plantations, to the slavery trade [1, 16]. In the first 100 years of colonization, the New-Christians reconquered the liberty lost in Iberia and particularly in the northeast of Brazil, the Jewish population increased to 14% of the total population [1, 16]. Brazilian New-Christians presented some interesting features that distinguish them clearly from the New-Christians who migrated to the countries of northern Europe or the Levant: the extensive admixture with the native population, creating deep roots in the new land and a full integration among the social and political organization [23, 24].

In the first quarter of the 17th century, the Portuguese Jews of Amsterdam created an extended merchant network that connected them with Portugal and Spain through their Marrano local centres, to the New world. If the war between Spain and Holland damaged the Sephardic commerce, the occupation of Recife by the *Dutch West India Company* gave it a new impetus and the first Jewish colony in South America, and also a congregation, (Tzur Israel), was established in 1638. In this period of prosperity, the Jewish population came to be one-third to half the total European civil population [23]. In the year of 1654 the Portuguese reconquered the Northeast and expelled the Dutch and as a consequence, the previous climate of tolerance ended abruptly and many Jews returned to Amsterdam or started new colonies in North America such as that of New Amsterdam (New York) [1, 16].

While in Brazil an Inquisition court comparable to the ones structured in Portugal (namely Lisbon, Coimbra and Évora) was never established, this territory was subordinate to the Court

of Lisbon. Jesuits were in Brazil the main agents of the Portuguese Inquisition. In the College of the “Companhia de Jesus” an Inquisitorial bureau was installed to execute the orders and accusations of the Inquisitors [23]. Between the 17th to the second half of the 18th centuries, 2169 people were suspected or sentenced for Judaizing by the Inquisition visitors [16].

Similarly to what happened in Portugal with the systematic persecutions and religious intolerance, also in Brazil the phenomenon of Crypto-Judaism emerged as a process to preserve a religious, ethnic and cultural identity. Believed to have vanished in the 16th – 17th centuries, Crypto-Judaism has arisen in the 20th-21st centuries, particularly in the Northeast Brazil, where some are claiming an unconditional return to Judaism because of the awareness of their Jewish or Marrano origin [24, 25].

Populations Genetics and Genetic Diversity

The fundamental aims of Population genetics are the comprehension of the evolutionary processes underlying the variability at a molecular level, between and among species` gene pools, which ultimately can elucidate their history, origin, phylogenetic relations, levels of genetic diversity and susceptibility to disease.

The development of Population Genetics can be traced back to the first decades of the 20th century, with the works of Ronald Fisher (1890–1962), JBS Haldane (1892–1964), and Sewall Wright (1889–1988), when Darwin´s ideas, particularly on natural selection, were combined with Mendelian genetics [26, 27]. They established the theoretical, mathematical foundations for what is now known as the modern *Synthetic Evolutionary Theory,* providing the formal keystones for natural selection measurements and also the statistical methods for assessing the effects of stochastic processes [28].

Because Mendel's genetic Principles of Inheritance were not known by Darwin, the initial Evolution Theory based on natural selection lacked specific mechanisms to explain the appearance of new variants [28]. The ability to understand this molecular genetic variation evolved in parallel with successive technological and scientific achievements, particularly in the second half of the 20th century: from Watson and Crick´s breaking of the genetic code and Smithies` development of starch-gel electrophoresis in the 1950s; the survey of human genetic diversity by Lewontin, M.Nei, Kimura and Crow in the 1960/70s and DNA sequencing methods described in the 1970s; the Development of Polymerase Chain Reaction (PCR) methods in the 1980s and the draft of human genome sequence in 2001, among many other accomplishments that could be referred to [29, 30].

At the end of the 1960s, Kimura [31, 32], and King and Jukes [33] questioned some of the assumptions of the Synthetic Theory of Evolution when analyzing amino acid substitution rates and allozyme polymorphisms and claimed that their genetic load would be too great if explained by selection alone. These findings culminated with the construction of the *Neutral Theory of Evolution* based on the premise that almost all mutations are selectively neutral and the majority of molecular variations result from genetic drift rather than natural selection [26, 28]. Since then, although surrounded by controversy from its appearance, the Neutral Theory has been widely

accepted in explaining molecular and genomic evolution [34, 35] but the debate on the role of natural selection still goes on today [36].

Michael Lynch [30], stated that “Nothing in evolution makes sense except in the light of population genetics”, showing the relevance of population genetics in the theoretical foundations for the comprehension of the evolutionary mechanisms. Evolution is the result of four different forces and Darwin proposed one of them, natural selection. The remaining mechanisms are considered nonadaptive, since they don´t rely on fitness properties.

Mutation and recombination are the key sources of variation, while genetic drift and gene flow change gene frequencies upon which selection will act [29, 30]. Population genetics pays special attention to the analysis of variation or maintenance of gene frequencies over generations and the processes behind those changes. A major contribution for this assessment is the Hardy–Weinberg Theorem that states that under particular conditions (absence of evolutionary forces and an infinite randomly sexually mating population), allelic and genotypic proportions of a population will stay in equilibrium over successive generations, allowing the estimation of allele and genotype frequencies from one generation to the next [29, 37]. Because of the imposed conditions, several factors can affect this equilibrium, such as the “effective population size”, meaning the breeding effective population, which is a fraction of the total population census, as generations do overlap. Fluctuations on the effective population size can impact the genetic diversity of a population through the processes of bottlenecks (a significant reduction of a previously large population with a consequent loss of diversity) and also through founder effects (a separation/migration of a subset of the total genetic diversity from an ancestral source population).

A real population is never “infinite” and large natural populations do not exhibit random mating. Instead, they might show inbreeding patterns (mating between people who are genetically related) or present positive or negative assortative mating (when mating pairs are more phenotypically similar or dissimilar respectively, than would be expected by random mating) [29, 37]. Inbreeding can increase the frequency of homozygotes and decrease the frequency of heterozygotes, but it does not directly change the allele frequencies, therefore it is not considered an evolutionary force [27].

The absence of equilibrium in a population can suggest the influence of the evolutionary forces referred to above. If a change in the allele frequencies is observed it represents the outcome of micro-evolutionary mechanisms. Mutation and recombination, as referred to before, are the two

main forces for the generation of genetic diversity. With recombination occurring through meiosis, the offspring can present new allelic combinations or haplotypes from those presented in the parental gene pool, thus increasing haplotype diversities. This process can be assessed by the study of whether a particular allele at different loci is associated with another allele, more or less often than would be expected by chance. This nonrandom association, Linkage Disequilibrium (LD), can indicate the action of evolutionary forces and contribute to the understanding of past evolutionary and demographic events, map genes associated with quantitative characters and inherited diseases, as well as comprehend the evolution of linked sets of genes [27, 38].

Mutation consists on the production of new alleles in the structure of chromosomes by random changes, since usually they do not follow any adaptive need of the organism. These changes can include point mutations, such as a single change of base pairs comprising transitions and transversions; insertions and deletions (INDELs) of one or more nucleotides in de DNA sequence; and changes that affect large portions of entire chromosomes, such as inversions and translocations. Mutations can happen at different sites along the genome and at a wide range of different rates. Mutations do modify allele frequencies since the ancestral state is chemically modified and tends to increase genetic diversity within and between populations. [27, 28, 37, 39]. Mutation alone might not produce major changes in allele frequencies; it requires other evolutionary forces to act upon the newly arising variants to increase or decrease its frequency in a gene pool [27, 39].

Gene flow, also known as admixture or migration, reflects the interbreeding between individuals from different populations. It can be “one-way” when a subset of a population admixes exclusively with another or “two-way” when the exchange of genes is made between two or more groups of people. Gene flow can affect variation both within and between populations since it can introduce new alleles into a population from elsewhere, increasing genetic diversity within that population. It also reduces variation between different populations by means of making the allele frequencies from the populations involved more similar over time [27, 29, 37].

Genetic drift is the random fluctuation of gene frequencies between generations. Since each generation represents a finite sample from a previous one, this allelic variation can result from a simple stochastic process of sampling. This phenomenon relates directly to the effective population size being sampled: the smaller the population is, the greater its effects will be and it is also more effective in isolated populations [39]. Genetic drift can increase, decrease, or

maintain the allele frequency in a population. Due to its random nature there is no inherent direction to this change. Eventually an allele is either lost (extinction) or replaces other alleles (fixation). The probability of extinction or fixation depends on the initial allele frequency and population size. Gene flow and genetic drift operate in opposition to each other. While genetic drift acts to make populations more differentiated, gene flow acts to make them more similar [27, 29].

The last of the evolutionary processes is natural selection. Darwin defended the survival of the fittest and mortality as the driving forces of evolution. The mechanisms of heredity and genetic variation were later added to the initial definition by Ronald Fisher in his new synthesis in evolution, the *Genetical Theory of Natural Selection.* Selection acts through differential mortality and fertility. It is not merely the survival of the individual but the successful reproduction and the transmission of the genes to the next generation in a particular environment. Fitness translates the genotypic reproductive success. It can act upon different biological levels including genes, larger sections of genetic material, organisms, populations' species or clades and in different life cycle stages, in viability and mortality, sexual selection, gamete selection and fecundity. There are diverse forms of selection, according to its effects on allele frequencies, such as positive or negative selection, when it favors the increase or the decrease of allele frequency respectively. As to its effects on genetic diversity, purifying selection tends to remove genetic variation and new alleles resulting from mutational processes while balancing selection roughly acts towards the maintenance of diversity [26, 29].

Interaction between the Forces Shaping Genetic Diversity

Overall, mutation, recombination and gene flow tend to increase genetic diversity while random genetic drift decreases it and finally, selection can do either, and all these forces interact dynamically.

As referred to above, genetic drift and gene flow operate in opposite ways, but over time these two forces can reach equilibrium, where differentiation among sub-populations remains constant. In population and evolutionary genetics this differentiation and genetic distance is measured by a descriptive statistic, the F_{ST}, (for fixation indices), and it compares the mean amount of genetic diversity due to allele frequency differences within sub-populations to that

found in the meta-population, in other words the correlation of randomly chosen alleles within the same sub-population relative to the entire population [29, 40]. F_{ST} is directly associated to the variance of allele frequencies among populations and the level of similarity among individuals within a population. Small F_{ST} values (close to zero) represent similar allele frequencies among sub-population and a high level of gene flow between them while a larger value (close to one), indicates different allele frequencies and highly differentiated sub-populations [40]. If natural selection favors a particular allele at a particular locus in some populations, F_{ST} values for that locus will present a higher value than that for other loci where differences among the sub-populations may simply result from genetic drift [40].

Other interacting forces include selection and mutation. The rate at which new mutations arise in a population is balanced by the subsequent elimination of each mutation by selection till equilibrium is reached [29, 37]. A similar process can happen through genetic drift, balancing the new variants created by mutation with those lost by drift [29]. Selection and genetic drift can be considered mutually exclusive since the first is systematic while the latter has a random nature and acts mostly on neutral alleles rather than those under strong selection pressure. Nevertheless, in very small sized populations, genetic drift can eventually act on genes that are not neutral, and the frequency of a disadvantageous allele can increase or even reach fixation [27, 37].

In natural populations each of these evolutionary forces does not represent independent or closed processes. Instead, they might act simultaneously over time throughout generations and their impact can be strongly influenced by characteristics such as the effective population size, degree of isolation or mating strategies. All this complexity of variables will ultimately create the genetic portrait of a particular population.

Assessing Genetic Diversity

Human genetic variation results originally from mutation and recombination which is then shaped by selection, mating systems, drift and gene flow. This variation or genetic diversity can be assessed using different types of genetic markers or polymorphisms in the DNA sequences. A polymorphism is the existence of two or more variants of DNA sequences or alternative alleles at a significant frequency, usually considered of ≥ 1% in human populations [41].

Historically, this variation was studied using the information provided by "classical" genetic markers such as proteins, blood groups and enzyme polymorphisms [42]. A series of methodological developments, namely rapid methods of DNA extraction, the use of restriction enzymes that cleaved DNA at specific sequences, DNA hybridization techniques, Polymerase Chain Reaction (PCR) and automated DNA sequencing, made possible the transition to molecular markers [37, 39]. These molecular markers made available the study of human variation directly at the level of the genetic material, addressing the physical bases of heredity [42]. Simultaneously to the methodological accomplishments, the number of detectable polymorphisms increased exponentially. This allowed a much more detailed perspective which moved towards the resolution of old anthropological questions such as the evolutionary relationship of humans to other species, the reconstruction of human migration patterns, and the assessment of historic relations between neighboring populations [43].

It has been more than a decade since the first global draft of the human genome was published [41, 44-46]. Since then, the once sketchy information on the biology and structure of the human genome began to emerge in an unprecedented way [47].

DNA markers can arise from virtually anywhere within the entire genome, whether the region is coding, regulatory, or noncoding. It is important nevertheless, when inferences on human demographic history are made based on DNA sequence data, to take into account the different features that affect diverse genomic compartments, such as the location and inheritance mode, as well as the factors shaping genetic diversity, such as the effective population size (N_e), recombination and mutation rates [48].

The human DNA is spread into 23 pairs of chromosomes, 22 pairs of autosomes and one pair of sex chromosomes (XX for females and XY for males). Besides this so called nuclear genome, there is also DNA outside the nucleus, the mitochondrial DNA (mtDNA) [43]. Each of these

genomic partitions has specific particularities. The usefulness of a particular polymorphism is profoundly associated with its origin and mutational rates and also the evolutionary process underlying its occurrence. Therefore, a specific type of molecular marker might provide deferential information and so be more or less suitable to answer particular population genetics or anthropological questions.

Most of our genome is noncoding (almost 99% of the total nuclear genome), or with unknown regulatory function, consequently these regions are not under selective constraints and therefore are more prone to variation. Molecular markers occurring in these parts of the genome are considered to be neutral and are most commonly used when demographic human history is surveyed, since they essentially reveal population level effects such as migration, admixture, drift and expansions [42, 43, 48, 49].

Although many other different kinds of molecular markers exist, in the context of this work, only Single Nucleotide Polymorphisms (SNPs), Short Tandem Repeats (STRs) and Insertion deletion polymorphisms (INDELs) will be considered.

DNA Genetic Markers

The most widespread types of sequence polymorphisms are the single nucleotide substitutions known as **SNPs (Single Nucleotide Polymorphisms)**. Most of the SNPs have just two alternative alleles (ancestral and mutant) in the DNA sequence and they represent rather rare mutational events, occurring at very low rates, ~$2.5x10^{-8}$ [50] in the human history. SNPs can appear throughout the nuclear and mitochondrial genomes in coding and noncoding or regulatory regions. Depending on the phenotypic effect, they can be classified as synonymous or non-synonymous, if the alleles encode the same or different products respectively, and silent or neutral if the SNP is located outside a coding region [51]. Although with very low rates, these are very common variants with an average of one SNP in each ~1,000 nucleotides between two randomly chosen chromosomes of a population [52]. If the minor allele frequency is more than 5% in a population, then it is considered “common SNP”; if the minor allele frequency is between 1% and 5%, then it is designated a “rare SNP”. The majority of heterozygous SNPs found in a particular individual are relatively common, while there are a far greater number of rare SNPs found at a population level [51, 52].

The first SNPs were detected with restriction enzymes (and were called RFLPs for restriction fragment length polymorphisms). With the development of genotyping techniques with high-throughput DNA sequencing technologies, such as microarrays that can assign thousands of SNPs simultaneously, the current number of known SNPs has increased remarkably. The dbSNP, a SNP database at NCBI, reports more than 97 million validated SNPs in the present Build 144.

The low mutation rate showed by SNPs translates a very low probability that they might occur recurrently during the evolution of modern humans, consequently individuals sharing the same SNP can be assumed to share a common ancestry, with alleles being identical by descent [43] rather than identical by state. As a result, these are the most appropriate markers in phylogenetic studies and migration patterns [53-57]; in population structure analysis [58-60]. They are also important in forensic studies when analyzing highly degraded samples, since they can be assessed using short PCR amplicons [61]; and also in relationship estimates [62].

Other prevailing genetic markers for population genetic analysis are the **Short Tandem Repeat** loci (**STRs**) or Short sequence repeats (SSRs), also known as microsatellites. Microsatellite DNA regions are a type of polymorphic variations in length due to a different number of short tandemly repeated units, usually of 2 to 6 base pairs [51, 52, 63]. These regions are scattered throughout the nuclear genome and can be found in exons, introns, regulatory regions as well as non-functional DNA sequences [51, 64].

STRs used as genetic markers are usually located in non-coding DNA, either in introns or intergenic sequences and therefore, assumed to evolve neutrally [65]. Contrarily to SNPs, which normally only have two alleles per locus, STRs are extremely polymorphic, with multiple alleles, often up to 20 or more at a particular microsatellite locus, and are also characterized by high levels of heterozygosity at a population level. These features make these sort of markers very useful for the discrimination between individuals at very close levels. Also, like SNPs, multiple microsatellite loci can be typed at the same time in a single PCR reaction, making these markers prone to multiplexing and automation and therefore particularly suitable for genome mapping, population genetic studies and forensic genetics [42, 51, 61, 64-66]. Although STRs mutation rates are not uniform among loci or among alleles, overall they are much higher than those observed for SNPs, ranging from 10^{-2} to 10^{-6} nucleotides per locus per generation [65, 67].

Microsatellites are assumed to primarily mutate by a mechanism of strand slippage during DNA replication, with the insertion or deletion of repeat units relative to the template strand. Even though a mismatch repair system exists for the correction of errors that may arise during DNA replication or recombination, such as an incorrect base or a small insertion–deletion loops incorporated into the new strand, a small fraction of these errors might not be repaired, ending up as microsatellite mutation events [63, 65].

While the fast mutation rate of microsatellites constitutes a major specificity for their broad use, on the other hand, the modeling of the mutational mechanism underlying its great genetic diversity may confound genetic inference and predictions [63]. Mutational models are used to extrapolate from the observed heterozygosity, the expected number of alleles in a population and also in the statistical analyses of genetic variation. Theoretical models such as the *Infinite Allele Model* or the *Infinite Sites Model* assume that each mutation randomly creates a new unique allele, and mutation at a given locus is expected to occur only once [63, 64].

While SNPs can be analyzed using these models, in STRs, because new alleles do not arise independently of the previous ones, these models are instead replaced by a more appropriate theoretical explanation for their evolutionary dynamics, the *Stepwise Mutation Model* [63, 65]. This model states that each mutational event results in a gain or a loss of a repeat unit. This suggests that (excluding homoplasies) two alleles differing by only one repeat motif are more closely related or share a more recent common ancestor than those diverging by numerous repeats [64]. Since microsatellite mutation rates vary across and within loci, the behavior of a particular locus might be better explained within a range from the *Infinite Allele Model* to the *Stepwise Mutation Model*, and thus all nuances affecting the mutation rate should be considered when genetic inferences are made based on these kind of data [59, 63, 68].

The last category of genetic markers assigned here are a length polymorphism created by the **insertion or deletion (INDELs)** of one or more nucleotides in the DNA sequence. These variants are ubiquitous in genomes and occur almost as frequently and with similar variation as SNPs, but with a multiplicity of sizes ranging from single base through much larger Indels. Some of them can have just two alleles (diallelic) while others might show multiple alleles (multiallelic) [69, 70]. An initial map of human INDEL variation reported 415,436 new INDELs spread throughout the human genome with approximately 36% of them located within the promoters, introns, and exons of known genes. Hence INDELs were thought to have some kind of influence in gene function and consequently in human traits and diseases [70, 71]. In contrast to the

prevalence and relative technical simplicity of SNPs, the focus on understanding the origin and functional effects of INDELs is a rather recent trend, probably associated with the difficulty in discovery and genotyping methods [72]. Over the last decades, significant advances, especially in next-generation sequencing technologies and computational variation analysis, have contributed largely to the discovery of new INDELs and their relevant contribution to genetic variation in humans as well as their influence on multiple human phenotypes [73].

Since the first large-scale efforts to identify these variants [74], the number of known INDELs have increased dramatically. A recent survey using population scale sequencing [72], reported a high-quality set of 1.6 million INDELS from 179 individuals representing 3 diverse human populations. INDEL variation was accessed with unprecedented resolution, showing that their mutation rates are highly heterogeneous across the genome, and that polymerase slippage is the mechanism underlying this heterogeneity [72].

INDELs can combine advantages from both SNPs and STRs, such as their widely spread distribution, their origin from a single mutation event which occurs at a low frequency and with low probability of presenting recurrent mutations, the significant allele frequency differences among distinct geographical population groups, and also the aptitude for large-scale multiplexing and automation. These features make INDELs a particularly useful genetic marker for studies of human identification and **ancestry informative markers (AIMs)** in forensic genetics and in population genetic studies, especially when bio-geographical ancestry analysis (in other words inferences on the ancestral origin of individuals) and proportions of admixture are relevant [75-77].

Regions of the Human Genome

When assaying patterns of human genetic diversity using several genetic markers, it is essential to have in mind the properties of the different compartments that constitute the human genome. The most relevant characteristics are summarized in the table below, and will be addressed in further detail in the next section:

Table 1. Comparison between different genomic compartments.

	Genomic compartment			
	Recombining Genome		Non-Recombing Genome	
Feature	**Autosomes**	**X chromosome**	**Y chromosome**	**mtDNA**
Location	Nuclear	Nuclear	Nuclear	Cytoplasmic
Inheritance	Biparental	Biparental*	Uniparental	Uniparental
Ploidy	Diploid	Haploid♂ Diploid♀	Haploid	Haploid
Relative *N*e	4	3	1	1
Diversity	High	Moderate	Low	Very High
Mutation rate**	Moderate	Low	High	Very High

*Heterogametic; in male sons the X chromosome (PARS excepted) is of uniparental origin.
** Values are variable according to different approaches, regions and type of genetic marker.
Adapted from: Daniel Garrigan and Michael F. Hammer (2006) [48] and Schaffner (2004) [78].

Recombining Genome

The Autosomes

Ninety-eight per cent of the human DNA is contained within the set of 22 pairs of autosomes [79], and so, the vast majority of genetic variation occurs in this genomic region. Autosomal polymorphisms are biparentally transmitted in a typical Mendelian mode of inheritance. At different loci, alleles do recombine and are assorted independently of one another when they are transmitted down each generation. This reshuffling generates new combinations of genetic material, as a result, they are not limited to either the paternal or the maternal side (as uniparental markers are), but rather can provide information about both parents' ancestries.

[43]. Thus, studying this biparentally inherited genome may have some advantages or disadvantages depending on the specific research question under consideration. Uniparental markers have been extensively used in the reconstruction of human evolutionary history and phylogenies [53], since they reflect a rather simple picture of the maternal and paternal genealogical past. In autosomes, as well as in the X chromosome, lineages are broken due to reshuffling in every generation, (with the exception of haplotype-block structures of the genome corresponding to regions of strong Linkage Disequilibrium - LD), which separates many different genealogical trees into different locations of the genome. On the other hand, the recombining genome can provide powerful data since it translates multiple genetic histories [80].

Information contained in autosomes can be considered overall to be neutral and unbiased when many independently inherited loci are being studied, since even if natural selection may be acting on particular loci, it is improbable that it acts in the same way in every loci. Recombination can also break down patterns of linkage disequilibrium and the study of these patterns and signals of decay may be used in the estimation of timings of admixture past events [79]. Autosomal loci are maintained in two copies in both males and females. As a result, when equal numbers of males and females are breeding, the effective population size (N_e) of the autosomes is the largest (4) compared to the X chromosomes (3), NRY (1) and mtDNA (1). So it is expected that this genomic region will show a deeper ancestry than the haploid systems, as for a neutrally evolving locus the expected Time of Most Recent Common Ancestor (TMRCA) is a function of the (N_e). Fluctuations in (N_e) together with studies of LD, are making the autosomes a precious tool for evolutionary inferences, particularly regarding periods of time in which no information is available from uniparental markers [48].

Autosomal diversity is also geographically structured, which makes this kind of data prone to the use of clustering algorithms, like STRUCTURE [81, 82]. This software uses multilocus genotypes to infer the structure of each population and to estimate admixture proportions, dividing global sample sets into continental clusters that reflect their origins, with no prior geographical or population affinity information [79] and has been used in several studies concerning Jewish populations [83-88].

The X Chromosome

The X chromosome is one of the two sex chromosomes (X and Y) in humans, which diverged from a pair of autosomes around 300 million years ago. Accumulated structural and functional variations changed them eventually into two different chromosomes. Successive rearrangements and mutation events, particularly on the Y chromosome, led to a progressive loss in recombination between the two homologous, despite their common autosomal origin, reducing the recombination process to the two small pseudoautosomal regions of the Y chromosome [78, 89]. Owing to different evolutionary paths, the X chromosome is present in two copies in females and just one copy in males and both sexes can transmit it (males only pass down their X chromosome to the daughters). This differential inheritance mode results in an unequal X-chromosome dosage between the XX female and XY male. Thus, a dosage compensation mechanism in which one of the two X chromosomes is inactivated in the female takes place [90]. Future remarks considering the X chromosome in the present work will designate the X-specific region that does not recombine with the Y chromosome.

Besides the differences in inheritance and ploidy, other specific characteristics make the X chromosome an increasingly popular genetic tool for population studies. For every three X chromosomes, only two of them recombine in a generation, thus mutations tend to be less frequent on the X chromosome than on autosomes because the nucleotide mutation rate in females is lower than in males. Since only 2/3 of X chromosomes recombine in each generation, the recombination rate for the X chromosome is practically two-thirds of the genome average. As a result, expected linkage disequilibrium is greater on the X chromosome and so regions with a single genetic history are expected to be larger than that of the autosomes. Considering the (N_e), the genetic diversity tends to be lower than in autosomes, as well as being a system which is more prone to the effects of genetic drift, more exposed to natural selection and population structure [78].

Due to all the above-mentioned features, the X chromosome is an ideal tool for understanding historical differences between males and females and the influence of natural selection [91], as well as differences in mutation rate and recombination patterns [92]. It detects sex biases on different timescales [93], sex specific behaviors, and in particular differences in migration rates between men and women [94]. It has also recently become a useful tool in forensic genetics for

Identity and Kinship testing [95, 96] and in population genetic history studies including ethnic groups such as the Roma [97, 98] and language isolates [99].

Non-Recombining Genome

The Y Chromosome

The Y chromosome comprises the biggest non-recombining block in the human genome and is also one of the most informative haplotyping systems used in a number of studies such as evolutionary population surveys, forensics, medical genetics, and genealogical reconstruction [53].

Like the X chromosome, the Y chromosome evolved from a pair of autosomes. In a first evolutionary step, it probably acquired a male-determining gene such as *SRY* followed by a progressive recombination suppression between the proto-sex chromosomes [100, 101]. The limitation in recombination between the emerging proto-X and proto-Y chromosomes was most likely the result of selection favoring genetic linkage between sexually antagonistic mutations (which are beneficial to one sex, but detrimental to the other), and the sex-determining region. In multiple successive steps, particular regions stopped the recombination process at distinct time points known as evolutionary strata. The oldest stratum in humans, which contains the *SRY,* dates from over 240 million years ago [100]. Consolidation of Y-haplotype linkage resulted from large scale inversions that disrupted alignment, and thus recombination, between increasingly larger regions of the X and Y chromosomes. Over the evolutionary timeframe, the mammalian X and Y started to diverge, with remarkably little structural changes in the X chromosome, whereas the Y rapidly began to degenerate [101].

The Y chromosome was thought to be a “genetic wasteland”, due to its degeneration by means of deletions and loss of gene function compared to the X chromosome [54]. The complete sequencing of the Non-Recombining region of Y chromosome [102] (NRY, also known as NRPY and MSY), changed this perspective significantly. Flanking the NRY are two small homolog regions, (PARs) that recombine during meiosis with the X chromosome corresponding to the pseudoautosomal regions. Further considerations regarding the Y chromosome in this work will denote exclusively the non-recombining portion. The NRY includes several blocks of heterochromatic sequence, and euchromatic regions that are comprise of the X-transposed

(containing only 2 genes), the X degenerate (with highly conservative genes), and the ampliconic sequences (arranged in palindrome structures, highly repetitive, with 9 gene families functionally specialized with testis-specific expression patterns) [102].

The lack of recombination and gene decay of the NRY with associated accumulation of deleterious mutations and a lower rate of adaptation compared to the X-linked genes, contributed largely to the idea that the Y chromosome was doomed to an eventual extinction [54, 100, 101, 103]. Recent studies however, have changed this perspective, providing new insights into evolutionary mechanisms, such as strong purifying selection and abundant gene conversion at palindromic sequences, which eventually justify its long-term survival [104-106].

Assuming a 1:1 sex ratio in a population, the (N_e) of the Y chromosome is ¼ of that of any autosome, one-third of that of the X chromosome and equal to that of mtDNA. Supposing that the same mutation processes act on all chromosomes, the diversity on the Y chromosome is thus expected to be lower than any other genomic region and also more susceptible to genetic drift [107]. This could result from neutral processes if the effective population size of males is reduced relative to females, but models of purifying selection and background selection have also been suggested to justify this low diversity [108].

If non-recombination could be evolutionarily interpreted as a weakness, in population genetic studies it surely can be regarded as an advantage. The Y chromosome (NRY) has a male-limited transmission and is inherited as a single block, so the mutations observed result exclusively from intra-allelic processes. Therefore, the haploid Y offers a proxy for studying stable male lineages throughout long time scales without the complexity due to reshuffling of the diploid systems [107]. By convention, Y chromosomes identified by SNPs are labeled to haplogroups and those defined by STRs are known as haplotypes. Data combining both SNPs and Y-STRs markers define lineages [53, 109]. The enormous amount of new polymorphisms defined in recent years has allowed the construction of a robust phylogenetic tree of human Y chromosomes that is constantly being updated [55, 110-113].

Due to its specific male inheritance and its particular haplogroup patterns with geographical distribution (**Figure 2**), the Y chromosome has been widely used to answer historical and anthropological questions regarding the Jewish people. Aspects such as their geographical origin, particular lineages such as the Cohanim or the Levites [114-117], relations between different ethnic groups such as Ashkenazim and Sephardim, as well as to non-Jewish host populations [118-123], in addition to medical issues, have been addressed lengthily [124, 125].

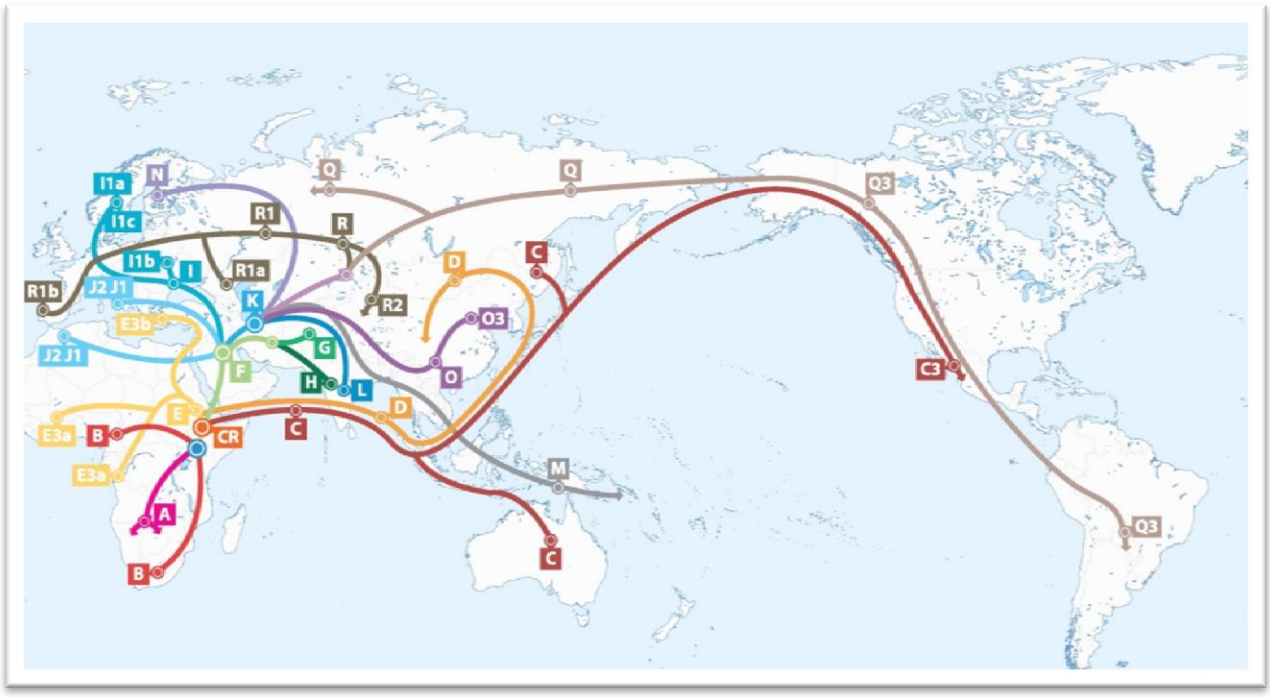


Figure 2. World distribution of the major human Y chromosome haplogroups. (adapted from http://www.familytreedna.com).

Mitochondrial DNA

Mitochondrial DNA is a circular, double-stranded molecule with a 16 569 base pair length, located outside the nucleus in the mitochondrion, a highly specialized organelle focused on the production of cellular energy through oxidative phosphorylation. The strands are classified by their nucleotide composition as heavy (H-strand – guanine rich), and as light strand (L-strand - cytosine-rich). Each cell contains several mitochondrion each one of them with many copies of mitochondrial genomes (from 100 to 10,000 copies per cell). The mitochondrial genome is comprised of two distinct regions, namely the coding and the non-coding regions. The coding region represents the majority of the human mtDNA and includes several genes,13 of which are essential components of the respiratory chain, 2 encoding ribosomal RNA (12S and 16S), and 22 transfer RNA (tRNA) genes spread between the protein-encoding genes [126, 127].

The non-coding region is a fragment of 1.1-kb length also known as D-loop, or control region, which is involved in the regulation of transcription and replication of the molecule. The D-loop

ranges from positions 16024 to 576 and comprehends the largest portion not directly associated with the synthesis of respiratory chain polypeptides or RNA [126, 127].

Similar to its uniparental male counterpart (NRY), the prevailing theory suggests that the mtDNA has an exclusive maternal inheritance pattern with no recombination [128] (a single case of paternal transmission in humans has been reported so far [129]). So, as it occurs with the NRY in the male-line, mtDNA is transmitted exclusively from the mother to the offspring unaltered, except for mutational events. This allows the definition of stable female related lineages back through time, tracing the maternal ancestry of a population. Similarly to the Y chromosome, mtDNA effective population size is ¼ relative to autosomes, and because of this lower effective size, mtDNA is more sensitive to demographic events such as bottlenecks, genetic drift and founder events.

Unlike nuclear DNA, mtDNA replication is not synchronized with the cell cycle and is constantly recycled. Each individual contain many copies of mtDNA and in theory their sequences are identical, a situation known as homoplasmy. However, as a result of inefficient mtDNA repair and a localized oxidative environment, as well as an increased replication, mtDNA mutations are very frequent. Mutations often co-exist with their wild-type counterpart, designated as heteroplasmy, in various proportions among different cell lineages. The processes underlying the transmission of the mutant mtDNA are not completely understood, but they comprehend a core of evolutionarily conserved mechanisms such as bottlenecks during germ cell development and strong purifying selection against specific mtDNA mutation types during maternal transmission [130]. The proportion of mutant variants has important consequences in understanding mitochondrial disease and evolution [131-133].

The mutation rate of mtDNA is known to have a much higher magnitude than the mutation rate in nuclear genes. However, the rates at which mutations occur or get fixed in mitochondria are not uniformly high along the molecule and its different functional domains. It is higher in the non-coding control region than in the coding region [131, 134]. Heterogeneous rates among sites associated with germ line and somatic heteroplasmies as well as the time dependence of the substitution rate, have made accurate estimates of human mtDNA mutation rate a tough challenge. Values are variable according to different approaches which can include pedigree, phylogenetic, phylogeographic and ancient DNA data, and also to the portion of mtDNA used, such as the whole molecule, the coding region, D-loop or smaller specific segments (for a review see Toomas Kivisild (2015)) [134].

The variability of mutation rates among different nucleotide positions and the identification of mutational hotspots can contribute significantly to distinguish if mtDNA molecules share nucleotides by descent or by state and thus allow a robust estimation of the mtDNA phylogeny [135].

Evolutionarily, common inherited mtDNA mutations, generally defined by SNPs, have created stable population subgroups defined by specific sequence variants known as haplogroups. These common, inherited, mtDNA variants are usually not heteroplasmic, and due to their selection neutrality have become fixed in the population. Variants found in a particular population are compared to the revised Cambridge Reference Sequence (rCRS) [136], a human mitochondrial complete genome sequenced from an European individual [137].

Since the first report of a complete mtDNA sequence, the number of available complete mitochondrial genomes has remarkably increased. These data allowed the construction of an updated comprehensive phylogenetic tree of global human mitochondrial DNA variation based on both coding and control region mutations [138]. rCRS is the reference for this phylogenetic tree. A more recent study suggested that an ancestral rather than a modern mitogenome from Europe, should serve as the center of the human mtDNA reference system, the Reconstructed Sapiens Reference Sequence (RSRS). In this phylogenetic tree, mtDNA mutations could be reported relative to their ancestral instead of to a derived status [139].

The special features of mtDNA, that is to say, the maternal inheritance, its high copy number, lack of recombination, and high mutation rate along with a specific pattern of geographical distribution of lineages (**Figure 3**), have made it the molecule of choice for studies of human population history and evolution. Among these studies, many have been focused on Jewish origins of maternal lineages, including Sephardim and Ashkenazim, as well on relations between Jewish and non-Jewish host populations [140-149].

Figure 3. World distribution of the major human mitochondrial DNA haplogroups. (adapted from http://www.familytreedna.com).

From Genetic Variation to Population History

Till very recently, when the history of a particular population was the focus of an investigation, the mathematical relationships underlying its genetic microevolution were far from being simply approached [27]. The usefulness of genetic markers for the analysis of anthropological questions of populational relationships in terms of geography and history has been evident since Cavalli-Sforza`s landmark work [150]. From a population genetics' perspective, history can be surveyed by looking at the record of genetic diversity. Very often, events in the past such as the migrations, population expansions and colonization occurring in the last few thousand years leave their signature in the genome [79]. Thus, the structure and history of populations can be accessed by interpreting genetic diversity in terms of the evolutionary forces that shaped their current genetic portrait. Cultural, linguistic, ethnic or religious differences between human groups, often studied in terms of culture contact, can also impact rates of gene flow and

fluctuations in population size due to ecological and demographic changes and might also affect levels of genetic drift [27].

Inferences about human history based on the information provided by genetic data relied initially on allele frequencies concluded directly from phenotypes. More recently, with the direct access to molecular DNA, genetic inferences rely on three main pathways: analysis of gene trees, analysis of summary statistics and coalescent-based analysis [29, 48].

The analysis of gene trees was primarily implemented for the non-recombining regions of the genome such as the NRY and mtDNA. Combining information about the ancestral/descendant relationships in the haplotype tree with the frequency and distribution of haplotypes in a sample of sequences, qualitative inferences can be made such as the place of the most recent common ancestor (PMRCA), for a particular genomic region, as well as the migration routes of derived haplotypes. Based on the assumptions of the molecular clock, in which the number of nucleotide mutations splitting two sequences is a linear function of time, the time to the most recent common ancestor (TMRCA) can also be inferred [29, 48].

A statistical summary approach does not embrace all the information contained in DNA sequence data although it allows comparisons between loci and populations. This approach can compare the observed values of the genetic data with that expected under a particular population genetics model. The expected values of many summary statistics are made under the standard Neutral Model which assumes a population in equilibrium. Summary statistics can obtain particular information from different aspects of DNA polymorphism data such as the estimated rates of recombination, gene flow, or be used to comprehend potential changes in effective population size. Interpreting summary statistics nevertheless should have in account that different evolutionary processes might result in identical values for one or more statistics [29, 48].

Coalescent-based methods require intense computational simulations to estimate statistical confidence compensating the drawbacks of summary statistic methods such as uncertainty of mutation rates and generation times and ambiguity about the demographic history of a population. This approach focuses on the coalescence of observed haplotypes in ancestral generations, from the present to the past, of only those chromosomes that appear in a sample, and so it is tremendously effective for simulating DNA sequence polymorphism data sets. However, a problem with model based methods is the bias due to incorrect specifications that

can affect the accuracy of the estimations and it is still difficult to include all complexities inherent to a real population within such models [29, 48].

Jewish populations have a myriad of specificities in their history. This includes a supposed common origin, a common religious and cultural identity background, besides the complexity behind the constant population movements with major Diasporas. All these particularities make the Jewish population a fascinating subject for population genetic studies, as mentioned above. Inferences on the levels and sex-biased admixture with host populations, founder effects along the Diaspora, gene flow directions and effects of genetic drift have been studied in these populations. While the analysis of their genetic variation can contribute greatly to the reconstruction of their history, inferences from genetic data also should be made with caution since all the complexity of the demographic and evolutionary processes acting to shape their extant genetic portrait is still a challenge, which needs to be accounted for.

CHAPTER II. AIMS

Aims

The main goal of this thesis was to contribute to the construction of a reference model for the history of Jewish communities of Portuguese origin, in which genetic and classical historical data interplay dynamically, in order to obtain insights into the ways the cultural bind (ethnicity/religion) is linked to biological kinship.

Special attention was primarily devoted to an extended characterization of the current genetic portrait of the Portuguese Jews, representatives of the original Sephardic group. Here the paternal and maternal lineages as well as autosomal and X chromosome linked markers were described. Uniparental markers analysis was then drawn out to the Diaspora groups available, such as those from Brazil and São Tome e Principe. As populations with complex demographic histories, a focus on the evolutionary processes underlying their actual genetic pools such as levels of admixture, gene flow or genetic drift, were assessed in order to amplify our knowledge of the far less studied Sephardic Jewish communities.

To achieve these main goals, specific questions were assigned, namely:

a) Were these communities at the time of the expelling decrees genetically distinct?

b) To what extent have the communities that stayed in Portugal kept not only their cultural identity, but also their genetic make-up?

c) Has the cultural isolation implied the impoverishment of genetic diversity?

d) Is the migrant communities' genetic profile consistent with a foundation by Iberian Jews or does it reveal the incorporation of other contributions?

e) Do the Diaspora communities show the same pattern or is it different, namely in terms of gender, as it is classical in colonial environments?

CHAPTER III. RESULTS

Research article:

Teixeira, J. C., Nogueiro, I., Goios, A., Gusmão, L., Amorim, A., & Alvarez, L. (2011). Mitochondrial DNA-control region sequence variation in the NE Portuguese Jewish community. *Forensic Science International: Genetics Supplement Series*, *3*(1), e51-e52.

G Model
FSIGSS-548; No. of Pages 2
ARTICLE IN PRESS

Forensic Science International: Genetics Supplement Series xxx (2011) xxx–xxx

Contents lists available at ScienceDirect

Forensic Science International: Genetics Supplement Series

journal homepage: www.elsevier.com/locate/FSIGSS

Mitochondrial DNA-control region sequence variation in the NE Portuguese Jewish community

J.C. Teixeira [a,b,*], I. Nogueiro [a,b], A. Goios [a], L. Gusmão [a], A. Amorim [a,b], L. Alvarez [a]

[a] IPATIMUP, Institute for Pathology and Molecular Immunology of the University of Porto, Porto, Portugal
[b] Faculty of Sciences of the University of Porto, Porto, Portugal

ARTICLE INFO

Article history:
Received 10 August 2011
Accepted 13 August 2011

Keywords:
mtDNA
Haplogroup frequencies
Crypto-Jews
Population substructure

ABSTRACT

The cultural phenomenon of Crypto-Judaism, defined as the secret adherence to Judaism while publicly professing another faith, arose in Portugal in the beginning of the 16th century after the Decree of Expulsion and the establishment of the Inquisition. Surprisingly, the scientific community acknowledged the persistence of Crypto-Judaic communities at the beginning of the 20th century in central and north-eastern regions of the country (e.g. Bragança). In the present work, we have sampled 56 unrelated individuals from the Bragança Jewish community aiming to characterize their maternal lineage. A 3348 bp mtDNA fragment was amplified and sequenced using mitochondrial-specific primers in order to obtain the entire control region. Haplogroup classification was performed according to current nomenclature. High frequencies were found for haplogroups H, HV0, T2, U2 and N1, indicating some degree of European admixture along with a remarkable signature of a Near East ancestry. These data confirm that the Crypto-Jews from Bragança were able to maintain not only their cultural identity but also some ancestral genetic identity, showing a significant population substructure within Portugal, with forensic relevance.

1. Introduction

Crypto-Judaism, defined as the secret adherence to Judaism while publicly professing another faith, arose in Portugal in the beginning of the 16th century after the Decree of Expulsion and the establishment of the Inquisition. Surprisingly, the scientific community acknowledged the persistence of Crypto-Judaic communities at the beginning of the 20th century in central and north-eastern regions of the country. A strong sense of a Jewish root is still well alive in several communities (e.g. Bragança), although the adhesion to Judaism is no longer a distinction mark. In this work, we aimed at characterizing the maternal lineages (mtDNA) of the referred communities and comparing them with other Jewish and non-Jewish populations.

2. Materials and methods

2.1. Population sampling and DNA analyses

Buccal cells collected on cytology brushes were obtained, under informed consent, from 56 unrelated self-designated Jewish males living in the Bragança district, Portugal. Personal inquiries were performed in order to avoid close kinship and confirm Jewish ancestry for at least three generations. DNA was extracted using Blood and Cell Culture DNA Spin kit (GENOMED, GmbH, Germany). The current population sample overlaps the one previously studied for the Y-chromosome [1]. For all samples, electropherograms were obtained from a 3348 bp fragment containing the entire mtDNA-control region as well as from three fragments encompassing coding region's relevant polymorphic positions for haplogroup (Hg) classification [2], previously amplified by polymerase chain reaction (PCR) using mtDNA specific primers [3].

2.2. Data analyses

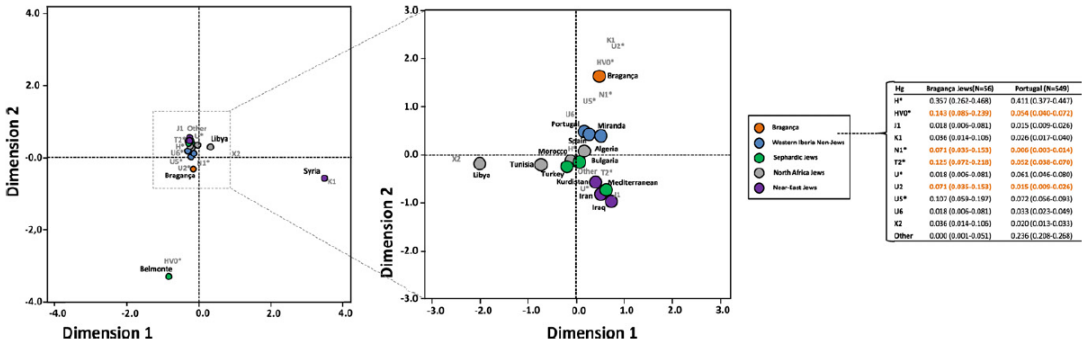
We compiled publicly available data on mtDNA Hgs' frequencies composition of 746 Jewish [4,5] and 749 Iberian-host individuals: 670 from Portugal [6; Quim Mairal, personal communication] and 79 from NW Spain [7]. Hg frequencies were determined by direct counting. In order to attain statistical significance for frequencies of the putative Jewish founding lineages, we calculated Bayesian 0.90 credible region (90% CR) using SAMPLING software (Vincent Macaulay, personal communication). Correspondence analysis using Hgs' frequencies was

* Corresponding author at: IPATIMUP, Institute of Pathology and Molecular Immunology of the University of Porto, Rua Dr. Roberto Frias s/n, 4200-465 Porto, Portugal. Tel.: +351 225570700; fax: +351 225570799.
E-mail address: jteixeira@ipatimup.pt (J.C. Teixeira).

1875-1768/$ – see front matter © 2011 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.fsigss.2011.08.025

Please cite this article in press as: J.C. Teixeira, et al., Mitochondrial DNA-control region sequence variation in the NE Portuguese Jewish community, Forensic Sci. Int. Gene. Suppl. (2011), doi:10.1016/j.fsigss.2011.08.025

G Model
FSIGSS-548; No. of Pages 2
ARTICLE IN PRESS



Hg	Bragança Jews(N=56)	Portugal (N=649)
H*	0.357 (0.162–0.468)	0.413 (0.377–0.447)
HV0*	0.143 (0.085–0.239)	0.054 (0.040–0.072)
J1	0.018 (0.006–0.081)	0.015 (0.009–0.026)
K1	0.036 (0.014–0.105)	0.026 (0.017–0.040)
N1*	0.071 (0.035–0.153)	0.006 (0.003–0.014)
T2*	0.125 (0.072–0.218)	0.052 (0.038–0.079)
U*	0.018 (0.006–0.081)	0.041 (0.046–0.080)
U3	0.071 (0.035–0.153)	0.016 (0.009–0.034)
U5*	0.107 (0.058–0.197)	0.073 (0.054–0.098)
U6	0.018 (0.006–0.081)	0.013 (0.023–0.049)
X2	0.036 (0.014–0.105)	0.020 (0.013–0.033)
Other	0.000 (0.001–0.051)	0.236 (0.208–0.268)

Fig. 1. Correspondence analysis plot illustrating the relationships among the Bragança Jewish community and the remaining considered North African, Near East and Sephardic Jewish populations, together with a general Portuguese and Spanish samples. Zoom plot corresponds to the same analysis excluding Syria and Belmonte populations, generated in order to increase resolution. Hg frequencies distribution in the Bragança Jewish community and in Portugal is also described, highlighting (orange) the putative Jewish founding lineages, with non-overlapping Bayesian 90% credible region.

performed and represented in a two-dimensional Euclidean space with SPPS ver. 15.0.1 (SPSS Inc.).

3. Results and discussion

In general, the Hgs' frequencies found in the Bragança Jewish community are in accordance with those expected for a typical Western European population [8], namely due to the frequency values obtained for Hgs H* (35.7%), U* (21.4%), J* (1.8%), T* (12.5) and X* (3.6%) (Fig. 1).

In order to identify traits of Jewish ancestry, we considered Hgs showing frequencies higher than 5% to be putative Jewish founding lineages, following the criterion established by Behar et al. [4]. As this criterion might be somehow arbitrary, a complementary step was adopted to address statistical significance, including only Hgs showing, as well, no overlapping Bayesian 0.90 credible region (90% CR) when compared to those found in the Portuguese-host population [6]. These criteria were met for Hgs HV0, N1, T2 and U2, pointing out the presence of several Jewish putative founding lineages.

The location of the sample in the correspondence analysis plots seems to be motivated by the high frequencies of the referred Hgs, which, together, account for 41% of the total variation (Fig. 1). These Hgs had been described to be of Near-East origin, with the exception of Hg HV0, which has not been yet clarified at a phylogeographic level, corroborating the strength of the combined criteria adopted.

Our data indicates some degree of population admixture, similarly to other diaspora Jewish populations. Therefore, the Hg frequencies' distribution seem to indicate population substructure within Portugal, which is highly relevant for forensic purposes and reveal the presence of a significant Near-East signature among the Bragança Jewish community.

Role of funding

This work was partially supported by Fundação para a Ciência e a Tecnologia through PhD grant SFRH/BD/73336/2010 and Post-doctoral grant SFRH/BPD/65000/2009.

Conflict of interest

None.

Acknowledgement

The authors would like to thank Cristina Santos from Universitat Autònoma de Barcelona for the most helpful discussion.

References

[1] I. Nogueiro, L. Manco, V. Gomes, et al., Phylogeographic analysis of paternal lineages in NE Portuguese Jewish communities, Am. J. Phys. Anthropol. 141 (March (3)) (2010) 373–381.
[2] M. van Oven, M. Kayser, Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation, Hum. Mutat. 30 (February (2)) (2009) E386–E394.
[3] A. Ramos, C. Santos, L. Alvarez, et al., Human mitochondrial DNA complete amplification and sequencing: a new validated primer set that prevents nuclear DNA sequences of mitochondrial origin co-amplification, Electrophoresis 30 (May (9)) (2009) 1587–1593.
[4] D.M. Behar, E. Metspalu, T. Kivisild, et al., Counting the founders: the matrilineal genetic ancestry of the Jewish Diaspora, PLoS One 3 (4) (2008) e2062.
[5] A. Picornell, P. Gimenez, J.A. Castro, et al., Mitochondrial DNA sequence variation in Jewish populations, Int. J. Legal Med. 120 (September (5)) (2006) 271–281.
[6] L. Pereira, M.J. Prata, A. Amorim, Diversity of mtDNA lineages in Portugal: not a genetic edge of European variation, Ann. Hum. Genet. 64 (November (Pt 6)) (2000) 491–506.
[7] O. Garcia, R. Fregel, J.M. Larruga, et al., Using mitochondrial DNA to test the hypothesis of a European post-glacial human recolonization from the Franco-Cantabrian refuge, Heredity 106 (January (1)) (2011) 37–45.
[8] E. Arroyo-Pardo, C. Baeza, E. Fernández, et al., Genetic history of the Iberian Peninsula, in: C. Santos, M. Lima (Eds.), Recent Advances in Molecular Biology and Evolution: Applications to Biological Anthropology, Research Singpost, Kerala, 2007, pp. 389–411.

Please cite this article in press as: J.C. Teixeira, et al., Mitochondrial DNA-control region sequence variation in the NE Portuguese Jewish community, Forensic Sci. Int. Gene. Suppl. (2011), doi:10.1016/j.fsigss.2011.08.025

Research article:

Nogueiro, I., Teixeira, J., Amorim, A., Gusmão, L., & Alvarez, L. (2014). Echoes from Sepharad: signatures on the maternal gene pool of crypto-Jewish descendants. *European Journal of Human Genetics*.

European Journal of Human Genetics (2014), 1–7

www.nature.com/ejhg

ARTICLE

Echoes from Sepharad: signatures on the maternal gene pool of crypto-Jewish descendants

Inês Nogueiro*[,1,2], João Teixeira[3], António Amorim[1,2], Leonor Gusmão[1,4] and Luis Alvarez[2]

The majority of genetic studies on Jewish populations have been focused on Ashkenazim, and genetic data from the Sephardic original source, the Iberian Peninsula, are particularly scarce. Regarding the mitochondrial genome, the available information is limited to a single Portuguese village, Belmonte, where just two different lineages (a single one corresponding to 93.3%) were found in 30 individuals. Aiming at disclosing the ancestral maternal background of the Portuguese Jewry, we enlarged the sampling to other crypto-Jewish descendants in the Bragança district (NE Portugal). Fifty-seven complete mtDNA genomes were newly sequenced and — in contrast with Belmonte — a high level of diversity was found, with five haplogroups (HV0b, N1, T2b11, T2e and U2e) being putatively identified as Sephardic founding lineages. Therefore — in sharp contrast with Belmonte — these communities have managed to escape the expected inbreeding effects caused by centuries of religious repression and have kept a significant proportion of the Sephardic founder gene pool. This deeper analysis of the surviving Sephardic maternal lineages allowed a much more comprehensive and detailed perspective on the origins and survival of the Sephardic genetic heritage. In line with previously published results on Sephardic paternal lineages, our findings also show a surprising resistance to the erosion of genetic diversity in the maternal lineages.
European Journal of Human Genetics advance online publication, 30 July 2014; doi:10.1038/ejhg.2014.140

INTRODUCTION

Sepharad, the traditional Jewish word for the Iberian Peninsula, became the name of the branch of the Jewish people who can trace their origin from there.[1] Their settlement in Iberia certainly occurred a long time ago, and the oldest archeological evidence found so far has a chronology of 390 CE (http://www.uni-jena.de/en/News/PM120525_Schrifttafel.html). Thus, the Jewish presence in Portugal precedes the nation's foundation in 1139 CE,[2] and several documents for the period 1279–1325 testify the presence of Jewish communes in the Bragança district, NE Portugal.[3] The degree of tolerance toward these communities was variable during the subsequent years and, in the sixteenth century, Iberian Jews were a demographically non-negligible minority with very heterogeneous social status. However, during that period (first in the United Kingdoms of Castile and Aragon under the rule of Catholic Kings and, shortly after, also in Portugal), Jews were forced into either conversion or expulsion. The Portuguese edict of expulsion was, however, far different from the one issued in Spain. As Jewish permanence in the country was intended, it was followed by several contradictory measures, as the forced baptism of 20 000 Jews from all over the country, who were prepared for exile and the forbiddance of inquiries on religion, during 20 years, for the *New-Christians*; hence, a consented crypto-Judaism (the secret adherence to Judaism while publicly professing another faith) was installed.[2,4,5] Although a decree to end the distinction between *Old* and *New-Christians* was issued in 1507, the inquisition was established in 1536 and the effective abolishment of the distinction would only happen three centuries later with the Pombaline law.[2]

As the inquisition mainly targeted crypto-Judaism, it was especially rampant in the most remote areas of Portugal, near the Spanish border, such as Bragança and Belmonte. The persecutions reached Bragança region, Trás-os-Montes, in 1582–1583 causing numerous accusations and arrests from the very beginning.[6–8] In the seventeenth and eighteenth centuries, the inquisitorial processes intensified and, as a result, there was a significant exodus to other countries, particularly of manufacturers and the merchant elite,[9] many of them preserving connections with their original communities for a long period.[7] The Jewish community of Bragança reappeared in the early twentieth century, gathering several families from the region, who had maintained their culture and religious secret practices for centuries. The estimated number of crypto-Jews at that time was around 700 to 800 people.[7] Although the community was dissolved shortly after, a strong sense of belonging is still well alive today among the Jewish descendants.

Contemporary Jewish communities have been genetically analyzed both from population genetics and medical perspectives by means of uniparental and recombining markers,[10–17] and more recently also through genome-wide approaches.[18–23] However, only a few reports have been published on Sephardic and crypto-Jewish descendants.[24–30]

Recent analysis of paternal lineages in Iberia points to a high proportion of Jewish ancestry.[10] Nonetheless, this work assumes an oversimplified parental population scenario and recognizes 'alternative possible sources for lineages ascribed a Sephardic Jewish origin'.

Concerning Portugal, little information exists. For the Y chromosome, samples from Belmonte were included in the study from Adams *et al*[10] but were pooled with non-Iberian Sephardic Jews; Nogueiro *et al*[27] found an unexpected high haplotype diversity for an isolated, small-size population, scattered over the Bragança district

[1]Population Genetics, Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP), Porto, Portugal; [2]Faculty of Sciences, Biology Department, University of Porto, Porto, Portugal; [3]Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany; [4]DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Rio de Janeiro, Brazil
*Correspondence: Dr I Nogueiro, Population Genetics, Ipatimup, Rua Dr Roberto Frias, s/n, Porto 4200-465, Portugal. Tel: +351225570700; Fax: +351225570799; E-mail: inogueiro@ipatimup.pt
Received 24 January 2014; revised 11 June 2014; accepted 20 June 2014

(NE Portugal). Concerning mtDNA, previous studies identified some lineages with signatures of Sephardic origins found in Jewish communities from Portugal (Belmonte), Bulgaria, Italy, Turkey, Yugoslavia, Greece, Netherlands, Surinam and Spain[12] and, more recently, South Texas, USA, and Mexico.[31,32] Nevertheless, data from the Iberian Peninsula, which constitutes the original geographic source of these populations, can be considered scarce: in Portugal, the information available until now was restricted to a single village, where just two different mtDNA lineages were found.[12] Here, we are deepening preliminary results on control region (CR)[33] to obtain a better picture of the Portuguese Jewish maternal lineages (Figure 1), in order to scrutinize whether the low diversity found in Belmonte is indeed a general hallmark of their mtDNA pool.

MATERIALS AND METHODS

Population sampling and DNA analyses

We sequenced complete mtDNA genomes from 57 unrelated self-designated Jews (and recognized as such by the host community) from the Bragança district, the same samples already characterized for Y chromosome[27] and the mtDNA CR.[33] Sampling criteria and collection method, as well as DNA extraction, are described in Nogueiro *et al.*[27] The study was approved by the Ethics Committee of the University of Porto (N°02/CEUP/2012) and appropriate informed consent was required from all subjects. Full mtDNA sequences were obtained using the protocol described in Ramos *et al.*[34,35] Sequences were aligned against the revised Cambridge Reference Sequence GenBank accession number NC_012920.1[36] using Genious software version 5.5.8 (http://www.geneious.com), variants were annotated following the HGVS rules (http://www.hgvs.org/mutnomen/). Haplogroups were classified following the updated mtDNA phylogeny, PhyloTree, mtDNA tree Build 16[37] (http://www.phylotree.org/) and assigned haplotypes were submitted to the EMPOP database[38] (http://empop.org/). The accession number for the sequences reported in this paper is EMP00619.

Data analysis

Based on the mtDNA haplotypes found, standard and molecular diversity indices were estimated using ARLEQUIN software v3.5.1.3.[39] For comparative purposes, original and publicly available data for mtDNA HVRI and HVRII (16 024–16 365 bp and 72–300 bp, respectively) were also compiled for 766 Jewish samples[12] and 884 Western Iberian non-Jews.[40–42]

Differences in haplogroup composition, defined according to CR polymorphisms (16 024–16 400 bp for HVRI, and 073–340 bp for HVRII), among the analyzed Jews from Bragança and a large sample of the Portuguese host population,[40] were assessed through Fisher's exact tests implemented in IBM SPSS software.[43]

Median-joining networks[44] (http://www.fluxus-engineering.com/sharenet.htm) of the (16 024–16 400 bp for HVRI, and 073–340 bp for HVRII) were constructed, for comparative purposes, using a compilation of original data on the Portuguese Jewish communities with publicly available data sets on the Portuguese host population,[40] Mirandese[41] and NW Spain.[41,42]

Complementarily Median-joining networks using complete sequences were initially constructed and drawn, and then — for better visualization — redrawn for the defined putative Jewish founding lineages together with available complete sequences from public databases NCBI-GeneBank (http://www.ncbi.nlm.nih.gov/genbank/), empop.org (http://empop.org/), mtDNA community (http://www.mtdnacommunity.org/) and from Zhang *et al*[45] work.

Figure 1 Geographic location of the sampling area in the Iberian Peninsula context. The zoomed area in the Iberian Peninsula represents the Portuguese district of Bragança and the Spanish province of Zamora (bold letter). Grey dots indicate the location of the sampling municipalities (uppercase) and villages (lowercase).

RESULTS AND DISCUSSION

Haplotype diversity

Using complete mtDNA sequences, we were able to identify 45 haplotypes and 32 different haplogroups; detailed information is shown in Supplementary Table 1. We estimated global haplotype and nucleotide diversity to be just slightly lower than in the Portuguese host population, and much higher than in the Portuguese Jewish community from Belmonte (Table 1). Theta *k* values (θk) were also calculated for both the complete mtDNA as well as for the CR (Table 1). As expected, θk obtained for the CR is lower than when calculated for the complete mtDNA sequences, emphasizing the importance of studying complete mtDNA sequences in order to obtain more precise information on the maternal lineage diversity, as well as a more accurate estimate of the female effective population size.

Estimated θk value for the total sample is similar to the one usually found in Sephardic Jews from other countries, but exceptionally high if compared with Belmonte (Table 1), and also similar to the non-Jews from the neighboring region of Miranda.[41] It is worth mentioning that the estimated number of putative female founders in each village of the Bragança district is higher than the current number of Jewish families. In fact, the samples included in the present work represent, if not the whole, the vast majority of the extant lineages, as increasing the sample size to reach predicted levels of saturation[46,47] would lead to the inclusion of closely related individuals.

The CR haplogroup frequencies of the Jews from Bragança together with the Portuguese host population and other geographical neighbors (Miranda[41] and NW Spain[41,42]) are presented in Supplementary Table 2. Differences in haplogroup composition were assessed through Fisher's exact test among the Jews from Bragança and the Portuguese populations[40] and Mairal *et al.*[41] We focused on Jewish haplogroups that showed higher frequencies when compared with the Portuguese population, as they are likely to represent Sephardic origins (in accordance with Behar *et al*).[12] Statistically significant differences were found for haplogroups HV0b (12.3% *vs* 0.0%; $P=0.000$), N1 (7.0% *vs* 1.3%; $P=0.014$), T2b11 (7.0% *vs* 0.0%; $P=0.000$), T2e (5.3% *vs* 0.7%; $P=0.019$) and U2e (7.0T *vs* 1.3%; $P=0.014$). These putative Jewish founding lineages account for 38.6% of the total sampled gene pool.

Sephardic lineages

The complex history of Jewish populations along the Mediterranean area is reflected on the haplogroup diversity observed in extant

Table 1 Diversity indices and neutrality test results calculated for the complete mtDNA haplotypes found in the samples from Bragança district Jews (considering the total sample as well as each village separately)

Population	N	K (% K)	S	*HD± SD*	$\pi \pm SD$	θk *(95% CI)*
Complete mtDNA haplotypes						
Portuguese Jews						
Total sample	57	45 (78.95)	286	0.987±0.007	0.002±0.001	96.747 (53.910;179.708)
Bragança	11	11 (100.00)	83	1.000±0.039	0.001±0.001	*
Argoselo	24	18 (75.00)	164	0.967±0.024	0.002±0.001	31.047 (13.714;73.620)
Carção	8	6 (75.00)	94	0.893±0.111	0.002±0.001	9.231 (2.613;34.341)
Vilarinho dos Galegos	13	11 (84.62)	108	0.974±0.039	0.001±0.001	30.896 (9.645;108.127)
Control region—HVRI (16 024–16 365 bp) and HVRII (72–300 bp)						
Portuguese Jews						
Total sample	57	35 (61.40)	61	0.967±0.012	0.014±0.008	37.616 (22.328;63.866)
Bragança	11	9 (81.82)	20	0.946±0.066	0.011±0.006	20.730 (6.297;73.767)
Argoselo	24	14 (58.33)	36	0.931±0.033	0.015±0.008	13.188 (6.121;28.635)
Carção	8	5 (62.50)	22	0.857±0.108	0.016±0.010	4.694 (1.394;16.076)
Vilarinho dos Galegos	13	11 (84.62)	27	0.974±0.039	0.012±0.007	30.896 (9.645;108.128)
Belmonte[12]	30	2 (6.67)	6	0.129±0.115	0.001±0.001	0.279 (0.065;1.097)
Other Sephardic and non-Ashkenazi Jews[12]						
Bulgaria	71	46 (64.79)	70	0.982±0.007	0.012±0.007	55.477 (34.470;90.159)
Turkey	123	85 (69.11)	109	0.985±0.005	0.013±0.007	120.394 (82.638;176.996)
Mediterranean	19	18 (94.74)	46	0.994±0.019	0.016±0.009	158.779 (40.578;673.927)
Algeria	20	15 (75.00)	45	0.968±0.029	0.013±0.007	25.594 (10.596;65.022)
Libya	83	36 (43.37)	63	0.922±0.019	0.013±0.007	23.631 (15.104;36.717)
Morocco	148	80 (54.05)	92	0.979±0.005	0.012±0.006	70.307 (50.484;97.941)
Tunisia	36	25 (69.44)	43	0.971±0.015	0.012±0.007	34.939 (18.044;69.325)
Kurdistan	12	9 (75.00)	28	0.939±0.063	0.013±0.007	14.686 (4.971;45.966)
Syria	4	3 (75.00)	13	0.833±0.210	0.012±0.009	3.766 (0.774;18.233)
Iran	82	43 (52.44)	76	0.971±0.008	0.016±0.008	35.821 (23.098;55.513)
Iraq	134	48 (35.82)	79	0.950±0.009	0.016±0.008	26.344 (18.151;37.920)
Western Iberia non-Jews[40–42]						
Portugal	549	353 (64.30)	171	0.986±0.001	0.013±0.007	425.966 (356.006;510.438)
Miranda	121	57 (47.11)	81	0.973±0.007	0.011±0.006	41.666 (28.672;59.784)
NW Spain	214	129 (60.28)	117	0.984±0.004	0.013±0.007	136.409 (103.182;180.727)

K: Number of different haplotypes; *S*: number of polymorphic sites; HD: haplotype diversity; π: nucleotide diversity averaged over loci; θ_K: theta estimator based on the number of different lineages; * cannot be computed when all haplotypes are different.

populations. An overview of their genetic composition is shown in Figure 2.

Until recently, haplogroup HV0, the ancestor of HV0b, was thought to have originated in Eastern Europe soon after the Last Glacial Maximum, having afterwards spread from there, following an east-west axis throughout Europe.[48] However, the low gene and nucleotide diversities found in this region and in Northern Africa, compared with the ones found in North-Central Europe, seem to exclude Eastern Europe as a possible focus of expansion.[49,50]

In the Iberian Peninsula, HV0 is a rare haplogroup, found at low frequencies in NW Spain, Zamora (4.7%),[51] a Spanish province geographically close to Bragança. Because of a lack of resolution, it is not possible to differentiate haplogroup V inside HV0 in the Portuguese population.[40] The same happens with the study of Mairal *et al*[41] focused on a linguistically isolated population, Miranda do Douro, from the same geographical region (NE Portugal). Nevertheless, and considering the whole HV0 branch (haplogroup V included), we observe a frequency of 14% for this haplogroup, higher than the frequencies found in the Portuguese and in the Miranda populations (5.2% and 8.3%, respectively). Moreover, although HV0 has very low frequencies in eastern European Jewish Ashkenazim,[52] we found a high prevalence of this lineage within the Jewish community of Bragança. Our results are in accordance with a previous report[12] describing HV0 as a Jewish founding lineage in Portugal, as in Belmonte 93.3% of the analyzed mtDNA genomes could be traced back to a single female, carrying an mtDNA within haplogroup HV0b. Considering solely the CR, in a comparison between samples from Bragança Jews (present work), Portugal,[40] Miranda[41] and NW Spain,[41,42] our sample does not share any haplotypes with the Portuguese population (Supplementary Figure 1).

The most parsimonious tree of HV0b sub-clade, including all available mitochondrial complete sequences (Supplementary Figure S2), shows that Bragança HV0b lineage shares a common private variant m.8520A>G with Belmonte.[12] Moreover, the Bragança samples cluster together, sharing a more recent variant not previously described m.10644G>A, which seems to have arisen locally. As no ethnic information is available for the remaining HV0b samples, and no introgression seems to have occurred with the Portuguese host-population, our results support the hypothesis that at least the HV0b-8520G haplotype is a Sephardic Jewish founding lineage.

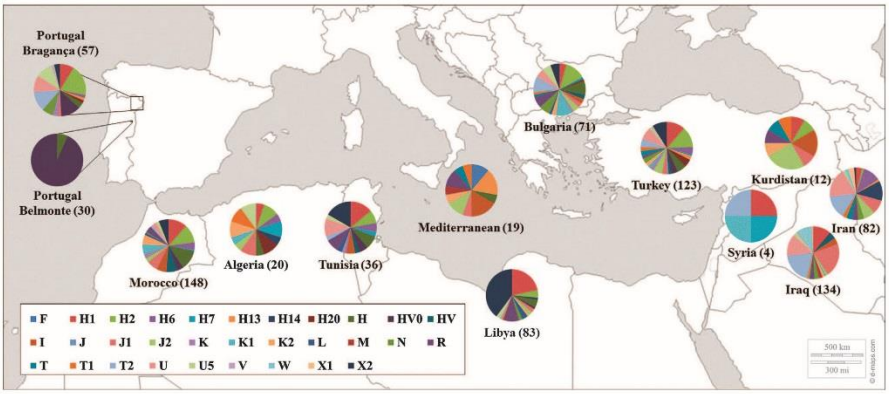


Figure 2 mtDNA haplogroup composition of the Sephardic Portuguese Jews (Bragança and Belmonte) and other Sephardic and non-Ashkenazi Jews. Haplogroups were defined for mtDNA HVRI and HVRII, 16024–16365 bp and 72–300 bp, respectively, according to Phylotree mtDNA tree Build 16 (other Sephardic and non-Ashkenazi Jews represent the same population as Table 1).

The high frequency of haplogroup N1 found within the Jewish descendants of Bragança (7%) differentiates this community from the general Portuguese population, which displays an average frequency of just 0.8%. Haplogroup N1, along with macrohaplogroups M and R, is one of the founder lineages of the Eurasian settlement around 50–63 kya[53] and comprises two major clades, N1a and N1b. N1a dates to approximately 20 kya and is a relatively rare haplogroup in Europe, reaching higher frequencies in Eastern Africa and in the Arabian Peninsula.[54] Although this haplogroup was found in Neolithic and Mesolithic skeletal remains from Central and Northwest Europe,[55–57] it was absent in previous studies on Iberian Neolithic populations, suggesting a rather uneven genetic and geographical spread of this lineage across Europe.[58,59] However, more recently, Gamba *et al*[60] detected this lineage in Neolithic northeast Iberia, evidencing a possible scenario of several Neolithic colonization events along the Mediterranean, from the Near East. The samples from Bragança cluster into sub-haplogroups N1a and N1b and similarly to what was observed for HV0, NJ networks of N1b revealed no CR haplotype sharing with the general Portuguese population, except for four individuals from Miranda, an isolated population from the same geographical area as well as with two samples from Zamora[41,42] (Supplementary Figure 3).

Within N1b, the N1b2 haplogroup is considered as a founding lineage in Ashkenazi Jews with a 'Hebrew/Levantine' origin.[61] However, an alternative European origin encompassing an assimilation of this lineage into the founding Ashkenazi population along the north Mediterranean coast has been recently proposed.[52] Unexpectedly, we found no samples belonging to the N1b2 haplogroup among the Bragança Jews — all N1b samples fall inside the N1b1 sister-clade (N1b1a2 and N1b1a5) and the N1a1a1a2 sub-haplogroup. According to the most parsimonious tree for complete sequences (Supplementary Figure 4), within sub-haplogroup N1b1a2, a transition at position m.204T>C defines a cluster with three samples, one from this study along with one from Zamora and another with an unknown origin. To further investigate this cluster, a search for similar CR haplotypes was performed in the EMPOP database and only three sequences were found. Two of them were from the same geographical area, namely Zamora (the same sample from Supplementary Figures 3 and 4) and Miranda do Douro,[41,42] both places with a well-documented history of Jewish presence. The remaining one was interestingly also from a Jew, an Ashkenazi from Hungary,[62] thus this transition could be identified as a Jewish, albeit not exclusively Sephardic founding lineage.

In the N1b1a5 sub-clade, the only matches found at EMPOP database, including the transition at position m.16311T>C, were with five samples from Miranda do Douro[41] and one from Uzbekistan,[63] about which no further information was available. As neither the Bragança Jews nor the Mirandese population share haplotypes with the Portuguese population, and given their geographic proximity in a rather remote and isolated area, as well as the fact that there was an organized Jewish community in Miranda, at least from the twelfth century on,[4] these individuals could easily also be Jewish descendants who lost memory of their origins or have not been detected as such, as in Mairal *et al*[41] sampling did not include ethnical criteria.

Regarding the N1a1a1a2 branch, the transitions at positions m.150C>T, m.4501C>T and m.11977C>T defines a cluster that includes the Bragança Jews (with no shared haplotypes with the Portuguese population). Lineages carrying the m.150C>T transition were not found in available databases and thus represent what could be a specific feature of the Sephardic Jews from NE Portugal.

Considering that haplogroup N1 is very rare in contemporary European populations, the low number of complete available sequences from the Near East, as well as the poor definition obtained so far for this haplogroup in the Iberian Neolithic samples[60] (N*), it seems at the moment hazardous to conclude whether the lineages found here are relics brought from the Near East by the first Jewish diasporas or were assimilated into the ancestral Iberian Jewish population in Europe

The T haplogroup in the Jewish population of Bragança is much more frequent than in Portugal[40] and pertains to the T2 branch, with two different sub-clades: T2b11 and T2e1. It has been proposed that haplogroup T, which diverged from the macro haplogroup JT around 40 kya, underwent a dramatic expansion from its geographical origin in the Near East into Europe before the Neolithic period.[64] An European indigenous dispersion has been proposed for T2 sub-clades, namely the T2b and T2e paraphyletic lineages.[64] Recently, several mitochondrial lineages, defined as the 'Neolithic

package' including T2, were also proposed to explain a rapid change from hunter-gatherering to farming, in the Early/Middle Neolithic, indicating a genetic flow from the Near East, Anatolia, and the Caucasus into Europe.[65] T2b is the most common lineage among T2, reaching higher frequencies in Western Europe than in the Near East.[64] Considering T2b branch, no haplotypes were shared between Jews and Portuguese individuals for the CR, demonstrating the absence of reciprocal introgression of T2b lineages between these two populations (Supplementary Figure 5).

Regarding the complete mtDNA genome, the Bragança samples nest in the branch defined by the transition at position m.9181A>G, along with a sample from Spain (Supplementary Figure 6). Interestingly, the Spanish sample is from the Zamora province (the same sample from Supplementary Figure 5) which, as stated above, is a geographically neighboring region to Bragança, at the Portuguese-Spanish border, and has also a well-documented Jewish past, with a well-established community dating back to 1259 CE.[66] Hence, the m.9181A>G variant within the T2b11 branch seems to be a regionally specific variant within the T2b11 branch, reflecting a Sephardic signature, given its absence from public databases. The additional motif m.4902A>G-m.8557G>A-m.16167C>T-m.16261C>T is only present among the Portuguese Jews.

Haplogroup T2e is spread throughout southern Europe and the Mediterranean, but is also found in Scandinavia, Egypt, the Near East and Arabia[64] and has been described as a founding lineage in the Bulgarian Sephardic community.[12] More recently, it was also found among Sephardic descendants in Turkey, in Northern Mexico and south Texas, USA, being interpreted as a 'Sephardic signature' inside haplogroup T.[31,32] According to the more recent nomenclature in PhyloTree built 16,[37] T2e's sub-branches, T2e1 and T2e1a, are defined by the variants m.41C>T and m.2308A>G, respectively. Considering the growing number of complete mtDNA sequences available at this time, it was possible to define a new sub-haplogroup T2e1a1, based on m.15499C>T variant. It is important to notice that variants at position m.41C>T are exceptionally infrequent along the mitochondrial phylogeny. Position m.41C>T seems quite unstable inside the new defined T2e1a1 sub-branch (Supplementary Figure 6). Considering a back-mutation at position m.41C>T and taking all mtDNA genomes available to date, the m.2308A>G-m.15499C>T-m.16114C>T-m.16192C>T motif defines one of the branches (T2e1a1a1) of Portuguese Sephardic signature within T2e, previously reported for a Sephardic sample[32] from Turkey, one of the Mediterranean countries that received exiled Iberian Jews. Two other samples shared this same motif, one from Mexico and another one from Texas. Without m.16114C>T, there was also a sample from Mexico. All these samples are from the study of Bedford *et al*[31] and although the ethnicity of these three samples is not known, their Iberian ancestry seems consistent with a Sephardic origin. The Bragança Jews present two further distinct variants in the Sephardic signature, m.13135G>A and m.7133C>T, the latter not described until now. In conclusion, the back-mutation at position 41 inside this new sub-haplogroup T2e1a1 is entirely associated, so far, with Sephardic or probable Sephardic ancestry.

Inside the T2e1 branch, a new sub-haplogroup T2e1b (Supplementary Figure 6), defined by the presence of variant m.9181A>G, has been proposed in PhyloTree built 16.[37] T2e1b is supported by 11 complete mtDNA sequences, including the Bragança Jews. Except for two samples from mtDNA Community database,[67] without information concerning their ethnicity, all the remaining nine individuals are Jews, Sephardim or Ashkenazim.

The last putative Jewish Sephardic founding lineage belongs to haplogroup U2, particularly its European U2e sub-haplogroup. This lineage is extremely rare in modern European populations and was recently found in Late Neolithic[68] and Iron Age populations, especially in the north of Europe.[69] It also appears with rather high frequencies in Croatia,[70] western Eurasia[71] and the Basque country.[50,72] All U2e individuals in this study belong to the U2e1a1 sub-clade. A high frequency of haplogroup U2 was found in the Bragança Jews (7.0%), whereas in Portugal it only reaches a frequency of 1.3%, being found absent in the Bragança district.[40] U2e1 lineages found in the Bragança Jewish population are shared with two individuals from Miranda; accordingly, this is the only putative Jewish founding lineage shared with the Portuguese population, when only the CR is looked upon (Supplementary Figure 7).

However, when considering complete mtDNA sequences, the Bragança clade is isolated from all other U2e1a1 haplotypes, except for another Jewish Ashkenazi sample from Moldova (Supplementary Figure 8), sharing the motif m.8014A>G—m.13708G>A, seemingly Jewish specific.

The strong founder effect previously reported in the maternal lineages of Portuguese Jews from Belmonte[12] cannot be considered a general trait of the Sephardic groups in Portugal. In fact, NE communities, despite preserving a distinctive lineage profile, displayed diversity levels similar to the host population. Remarkably, for two of the founder lineages (T2e1b and U2e1a), defined by the complete mitochondrial genome, the shared sequences belong to both Sephardic as well as Ashkenazi Jews. Two possible scenarios could accommodate this finding: either the defining variants for each branch could have arisen before the separation between the two Jewish groups; or there may have been recent introgression of Sephardic lineages into Ashkenazim communities in the north of Europe. More complete sampling and complete sequences will contribute in the clarification of which one is more likely. In any case, it must be said that although not frequent, marriages between the two communities occurred (especially) in the sixteenth and seventeenth centuries, namely among the elite sugar traders, with the descendants assimilated into the Ashkenazi community.[73,74]

The estimates of both the diversity levels and the number of female effective-population founders point at a stable size of the studied populations, in agreement with previous findings for the male counterparts.[27] As expected, it was possible to identify some Sephardic signatures as well as signs of introgression from the host non-Jewish population. This gene flow seems to have been mutual as the maintenance of the observed diversity levels can only be explained by a number of founders that is higher than the Jewish families, which can still be traced today.

A study of recombinant markers is now required to unveil the reproductive strategies that have sustained this ancestral signature along with a moderate degree of admixture with the host population.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Walther Parson and Mannis van Oven for the helpful discussion and suggestions in the phylogenetic analysis, Doron Behar and Felice L Bedford for the additional information concerning Jewish samples. Elvira Mea, Jorge Martins, Fernanda Guimarães, Jits van Straten and Harmen Snel for their contribution on Jewish history. Kevin Charles Rowe for the English language revision and Sofia Marques for the aid in the preparation of figures. All DNA donors for their collaboration and generosity. This work was supported by

FCT, Portuguese Foundation for Science and Technology, financed by the European Social Funds (COMPETE-FEDER), project PTDC/ATP-DEM/4545/2012. LA (SFRH/BPD/65000/2009), IN (SFRH/BD/73336/2010) and JT (SFRH/BD/77043/2011) are supported by FCT fellowships and funded by POPH-QREN—Promotion of scientific employment, supported by the European Social Fund and national funds of the Ministry of Education and Science. IPATIMUP is an Associate Laboratory of the Portuguese Ministry of Science, Technology and Higher Education and is partially supported by FCT.

1 Tavares MJPF: *A herança judaica em Portugal*. Lisboa: CTT Correios de Portugal, 2004.
2 Martins J: *Portugal e os judeus: Judaísmo e anti-semitismo no século XX*. Lisboa: Vega, 2006.
3 Garcia MA: *Os judeus de Belmonte: os caminhos da memória*. Lisboa: Instituto de Sociologia e Etnologia das Religiões, Universidade Nova de Lisboa, 1993.
4 Pignatelli M: *A comunidade israelita de Lisboa: o passado e o presente na construcção da etnicidade dos judeus de Lisboa*. Lisboa: Universidade Técnica de Lisboa, Instituto Superior de Ciências Sociais e Políticas, 2000.
5 Saraiva AJ, Salomon HP, Sassoon ISD: *The Marrano Factory: The Portuguese Inquisition and Its New Christians 1536–1765*. Leiden: Brill, 2001.
6 Azevedo EM: Cristãos Novos No Nordeste Transmontano: *Los caminos de Cervantes y Sefarad: actas del II Congreso Internacional*. Zamora, Bragança, Puebla de Sanabria, Benavente: Asociación Caminos de Cervantes y Sefarad, 1995.
7 Mea E: O Judaísmo no Séc. XX—A fénix renascida em Trás-os-Montes: *Congresso Histórico 450 Anos da Fundação da Diocese de Bragança-Miranda*. Bragança: Fundação da Diocese de Bragança-Miranda, 1996.
8 Mea EA: *A inquisição de Coimbra no século XVI: a instituição, os homens e a sociedade*. Porto: Fundação Engo Antônio de Almeida, 1997.
9 Paulo A: *Os judeus secretos em Portugal*. Fafe: Editorial Labirinto, 1985.
10 Adams SM, Bosch E, Balaresque PL *et al*: The genetic legacy of religious diversity and intolerance: paternal lineages of Christians, Jews, and Muslims in the Iberian Peninsula. *Am J Hum Genet* 2008; **83**: 725–736.
11 Bauchet M, McEvoy B, Pearson LN *et al*: Measuring European population stratification with microarray genotype data. *Am J Hum Genet* 2007; **80**: 948–956.
12 Behar DM, Metspalu E, Kivisild T *et al*: Counting the founders: the matrilineal genetic ancestry of the Jewish Diaspora. *PLoS One* 2008; **3**: e2062.
13 Hammer MF, Redd AJ, Wood ET *et al*: Jewish and Middle Eastern non-Jewish populations share a common pool of Y-chromosome biallelic haplotypes. *Proc Natl Acad Sci USA* 2000; **97**: 6769–6774.
14 Kopelman NM, Stone L, Wang C *et al*: Genomic microsatellites identify shared Jewish ancestry intermediate between Middle Eastern and European populations. *BMC Genet* 2009; **10**: 80.
15 Olshen AB, Gold B, Lohmueller KE *et al*: Analysis of genetic variation in Ashkenazi Jews by high density SNP genotyping. *BMC Genet* 2008; **9**: 14.
16 Ostrer H: A genetic profile of contemporary Jewish populations. *Nat Rev* 2001; **2**: 891–898.
17 Price AL, Butler J, Patterson N *et al*: Discerning the ancestry of European Americans in genetic association studies. *PLoS Genet* 2008; **4**: e236.
18 Seldin MF, Shigeta R, Villoslada P *et al*: European population substructure: clustering of northern and southern populations. *PLoS Genet* 2006; **2**: e143.
19 Campbell CL, Palamara PF, Dubrovsky M *et al*: North African Jewish and non-Jewish populations form distinctive, orthogonal clusters. *Proc Natl Acad Sci USA* 2012; **109**: 13865–13870.
20 Ostrer H, Skorecki K: The population genetics of the Jewish people. *Hum Genet* 2013; **132**: 119–127.
21 Atzmon G, Hao L, Pe'er I *et al*: Abraham's children in the genome era: major Jewish diaspora populations comprise distinct genetic clusters with shared Middle Eastern Ancestry. *Am J Hum Genet* 2010; **86**: 850–859.
22 Behar DM, Yunusbayev B, Metspalu M *et al*: The genome-wide structure of the Jewish people. *Nature* 2010; **466**: 238–242.
23 Velez C, Palamara PF, Guevara-Aguirre J *et al*: The impact of Converso Jews on the genomes of modern Latin Americans. *Hum Genet* 2012; **131**: 251–263.
24 Gerber S, Rozet JM, Takezawa SI *et al*: The photoreceptor cell-specific nuclear receptor gene (PNR) accounts for retinitis pigmentosa in the Crypto-Jews from Portugal (Marranos), survivors from the Spanish Inquisition. *Hum Genet* 2000; **107**: 276–284.
25 Sutton WK, Knight A, Underhill PA, Neulander JS, Disotell TR, Mountain JL: Toward resolution of the debate regarding purported crypto-Jews in a Spanish-American population: evidence from the Y chromosome. *Ann Hum Biol* 2006; **33**: 100–111.
26 Bordenave K, Griffith J, Hordes SM, Williams TM, Padilla RS: The historical and geomedical immunogenetics of pemphigus among the descendants of Sephardic Jews in New Mexico. *Arch Dermatol* 2001; **137**: 825–826.
27 Nogueiro I, Manco L, Gomes V, Amorim A, Gusmao L: Phylogeographic analysis of paternal lineages in NE Portuguese Jewish communities. *Am J Phys Anthropol* 2010; **141**: 373–381.
28 Picornell A, Gomez-Barbeito L, Tomas C, Castro JA, Ramon MM, Mitochondrial DNA: HVRI variation in Balearic populations. *Am J Phys Anthropol* 2005; **128**: 119–130.
29 Tomas C, Picornell A, Castro JA *et al*: Genetic variability at nine STR loci in the Chueta (Majorcan Jews) and the Balearic populations investigated by a single multiplex reaction. *Int J Legal Med* 2000; **113**: 263–267.
30 Tomas C, Picornell A, Castro JA, Ramon MM: Genetic analyses in five western Mediterranean populations: variation at five tetrameric short tandem repeat loci. *Hum Biol* 2001; **73**: 349–363.
31 Bedford FL, Yacobi D, Felix G, Garza FM: Clarifying Mitochondrial DNA Subclades of T2e from Mideast to Mexico. *J Phylogenetics Evol Biol* 2013; **1**: 121.
32 Bedford FL: Sephardic signature in haplogroup T mitochondrial DNA. *Eur J Hum Genet* 2012; **20**: 441–448.
33 Teixeira JC, Nogueiro I, Goios A, Gusmão L, Amorim A, Alvarez L: Mitochondrial DNA-control region sequence variation in the NE Portuguese Jewish community. *FSI Genet Suppl Ser* 2011; **3**: e51–e52.
34 Ramos A, Santos C, Alvarez L, Nogues R, Aluja MP: Human mitochondrial DNA complete amplification and sequencing: a new validated primer set that prevents nuclear DNA sequences of mitochondrial origin co-amplification. *Electrophoresis* 2009; **30**: 1587–1593.
35 Ramos A, Santos C, Barbena E *et al*: Validated primer set that prevents nuclear DNA sequences of mitochondrial origin co-amplification: a revision based on the New Human Genome Reference Sequence (GRCh37). *Electrophoresis* 2011; **32**: 782–783.
36 Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N: Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 1999; **23**: 147.
37 van Oven M, Kayser M: Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* 2009; **30**: E386–E394.
38 Parson W, Dur A: EMPOP–a forensic mtDNA database. *Forensic Sci Int Genet* 2007; **1**: 88–92.
39 Excoffier L, Lischer HE: Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 2010; **10**: 564–567.
40 Pereira L, Prata MJ, Amorim A: Diversity of mtDNA lineages in Portugal: not a genetic edge of European variation. *Ann Hum Genet* 2000; **64**: 491–506.
41 Mairal Q, Santos C, Silva M *et al*: Linguistic isolates in Portugal: insights from the mitochondrial DNA pattern. *Forensic Sci Int Genet* 2013; **7**: 618–623.
42 Ramos A, Santos C, Mateiu L *et al*: Frequency and pattern of heteroplasmy in the complete human mitochondrial genome. *PLoS One* 2013; **8**: e74636.
43 IBM Corp. Released *IBM SPSS Statistics for Windows, Version 21.0*. Armonk, NY: IBM Corp,, 2012.
44 Bandelt HJ, Forster P, Sykes BC, Richards MB: Mitochondrial portraits of human populations using median networks. *Genetics* 1995; **141**: 743–753.
45 Zhang C, Huang VH, Simon M *et al*: Heteroplasmic mutations of the mitochondrial genome cause paradoxical effects on mitochondrial functions. *FASEB J* 2012; **26**: 4914–4924.
46 Helgason A, Sigureth ardottir S, Gulcher JR, Ward R, Stefansson K: mtDNA and the origin of the Icelanders: deciphering signals of recent population history. *Am J Hum Genet* 2000; **66**: 999–1016.
47 Pereira L, Cunha C, Amorim A: Predicting sampling saturation of mtDNA haplotypes: an application to an enlarged Portuguese database. *Int J Legal Med* 2004; **118**: 132–136.
48 Torroni A, Bandelt HJ, Macaulay V *et al*: A signal, from human mtDNA, of postglacial recolonization in Europe. *Am J Hum Genet* 2001; **69**: 844–852.
49 Garcia O, Fregel R, Larruga JM *et al*: Using mitochondrial DNA to test the hypothesis of a European post-glacial human recolonization from the Franco-Cantabrian refuge. *Heredity (Edinb)* 2011; **106**: 37–45.
50 Cardoso S, Alfonso-Sanchez MA, Valverde L *et al*: The maternal legacy of Basques in northern navarre: New insights into the mitochondrial DNA diversity of the Franco-Cantabrian area. *Am J Phys Anthropol* 2011; **145**: 480–488.
51 Alvarez L, Santos C, Ramos A *et al*: Mitochondrial DNA patterns in the Iberian Northern plateau: Population dynamics and substructure of the Zamora province. *Am J Phys Anthropol* 2010; **142**: 531–539.
52 Costa MD, Pereira JB, Pala M *et al*: A substantial prehistoric European ancestry amongst Ashkenazi maternal lineages. *Nat Commun* 2013; **4**: 2543.
53 Macaulay V, Hill C, Achilli A *et al*: Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science* 2005; **308**: 1034–1036.
54 Fernandes V, Alshamali F, Alves M *et al*: The Arabian cradle: mitochondrial relicts of the first steps along the southern route out of Africa. *Am J Hum Genet* 2012; **90**: 347–355.
55 Haak W, Balanovsky O, Sanchez JJ *et al*: Ancient DNA from European early neolithic farmers reveals their near eastern affinities. *PLoS Biol* 2010; **8**: e1000536.
56 Palanichamy MG, Zhang CL, Mitra B *et al*: Mitochondrial haplogroup N1a phylogeography, with implication to the origin of European farmers. *BMC Evol Biol* 2010; **10**: 304.
57 Deguilloux MF, Soler L, Pemonge MH, Scarre C, Joussaume R, Laporte L: News from the west: ancient DNA from a French megalithic burial chamber. *Am J Phys Anthropol* 2011; **144**: 108–118.
58 Sampietro ML, Lao O, Caramelli D *et al*: Palaeogenetic evidence supports a dual model of Neolithic spreading into Europe. *Proc R Soc B* 2007; **274**: 2161–2167.
59 Hervella M, Izagirre N, Alonso S *et al*: Ancient DNA from hunter-gatherer and farmer groups from Northern Spain supports a random dispersion model for the Neolithic expansion into Europe. *PLoS One* 2012; **7**: e34417.

60 Gamba C, Fernandez E, Tirado M *et al*: Ancient DNA from an Early Neolithic Iberian population supports a pioneer colonization by first farmers. *Mol Ecol* 2012; **21**: 45–56.
61 Behar DM, Metspalu E, Kivisild T *et al*: The matrilineal ancestry of Ashkenazi Jewry: portrait of a recent founder event. *Am J Hum Genet* 2006; **78**: 487–497.
62 Brandstatter A, Egyed B, Zimmermann B, Tordai A, Padar Z, Parson W: Mitochondrial DNA control region variation in Ashkenazi Jews from Hungary. *Forensic Sci Int Genet* 2008; **2**: e4–e6.
63 Irwin JA, Ikramov A, Saunier J *et al*: The mtDNA composition of Uzbekistan: a microcosm of Central Asian patterns. *Int J Legal Med* 2010; **124**: 195–204.
64 Pala M, Olivieri A, Achilli A *et al*: Mitochondrial DNA signals of late glacial recolonization of Europe from near eastern refugia. *Am J Hum Genet* 2012; **90**: 915–924.
65 Brandt G, Haak W, Adler CJ *et al*: Ancient DNA reveals key stages in the formation of central European mitochondrial genetic diversity. *Science* 2013; **342**: 257–261.
66 Casar MFG: *El pasado judío de Zamora*. Valladolid: Junta de Castilla y León, Consejería de Cultura y Turismo, 1992.
67 Behar DM, van Oven M, Rosset S *et al*: A 'copernican' reassessment of the human mitochondrial DNA tree from its root. *Am J Hum Genet* 2012; **90**: 675–684.
68 Lee EJ, Makarewicz C, Renneberg R *et al*: Emerging genetic patterns of the european neolithic: perspectives from a late neolithic bell beaker burial site in Germany. *Am J Phys Anthropol* 2012; **148**: 571–579.
69 Melchior L, Gilbert MT, Kivisild T, Lynnerup N, Dissing J: Rare mtDNA haplogroups and genetic differences in rich and poor Danish Iron-Age villages. *Am J Phys Anthropol* 2008; **135**: 206–215.
70 Jeran N, Havas Augustin D, Grahovac B *et al*: Mitochondrial DNA heritage of Cres Islanders–example of Croatian genetic outliers. *Coll Antropol* 2009; **33**: 1323–1328.
71 Quintana-Murci L, Chaix R, Wells RS *et al*: Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. *Am J Hum Genet* 2004; **74**: 827–845.
72 Cardoso S, Villanueva-Millan MJ, Valverde L *et al*: Mitochondrial DNA control region variation in an autochthonous Basque population sample from the Basque Country. *Forensic Sci Int Genet* 2012; **6**: e106–e108.
73 Roth C, Spivak A: *Historia de los marranos*. Editorial. Israel 1946.
74 Mea E, Problemática do Judaísmo. A: Séculos XVI-XVII: *Congresso Internacional Inquisição Portuguesa. Tempo, Razão e Circunstância*. Lisboa, S Paulo, 2007.

Supplementary Information accompanies this paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)

Echoes from Sepharad: signatures on the maternal gene pool of crypto-Jewish descendants – Supplementary Material

Table S1. List of complete mtDNA haplotypes and corresponding haplogroups for the 57 Jewish samples included in this study from Bragança District

Region: Argozelo
Longitude/Latitude: **41.38 / 6.36**

Sample	Haplogroup	Haplotype
A01	H6a1a1a	239C,263G,315.1C,750G,1438G,3915A,4727G,4769G,4991A,5460A,7325G,8860G,9362G,9380A,11253C,11611A,15326G, 16311C,16362C, 16482G
A02	HV0b	72C,195C,198T,263G,315.1C,750G,1438G,2706G,4769G,7028T,8520G,8860G,10644A,15326G,16298C
A03	T2b11	73G,207A,263G,315.1C,709A,750G,930A,1438G,1888A,2706G,3398C,4216C,4769G,4902G,4917G,5147A,7028T,8557A, 8697A,8860G,9181G,10463C,11251G,11719A,11812G,13368A,14233G,14766T,14905A,15326G,15452A,15607G,15928A, 16126C,16167T,16261T,16294T,16296T,16304C,16519C
A04	U5a1a1	73G,263G,315.1C,750G,1438G,1700C,1819C,2706G,3197C,4769G,5495C,5806C,7028T,7670G,8860G,9477A,11467G, 11719A,12308G,12372A,13617C,14766T,14793G,15218G,15326G,15924G,16256T,16270T,16399G
A05	U5a1a1 +152	73G,152C,263G,315.1C,750G,1438G,1700C,1819C,2706G,3197C,4769G,5495C,5806C,7028T,7670G,8860G,9477A,11467G, 11719A,12308G,12372A,13617C,14766T,14793G,15218G,15326G,15924G,16256T,16270T,16399G

A06	H5b1	146C,195C,263G,309.1C,315.1C,456T,499A,750G,1438G,4769G,5471A,8860G,14497G,15326G,16304C
A07	HV0b	72C,195C,198T,263G,315.1C,750G,1438G,2706G,4769G,7028T,8520G,8860G,10644A,15326G,15951G,16298C
A08	T2e	73G,150T,263G,309.1C,315.1C,709A,750G,1438G,1888A,2308G,2706G,4216C,4769G,4917G,7028T,8697A,8860G,10463C,11251G,11719A,11812G,13135A,13368A,14233G,14766T,14905A,15326G,15452A,15499T,15607G,15928A,16114T,16126C,16153A,16192T,16294T,16519C
A09	H3v	263G,315.1C,408A,750G,1438G,4769G,6776C,8706G,8860G,9055A,9698C,10398G,10550G,14148G,15326G,16519C
A10	T2b11	73G,207A,263G,315.1C,709A,750G,930A,1438G,1888A,2706G,3398C,4216C,4769G,4902G,4917G,5147A,7028T,8557A,8697A,8860G,9181G,10463C,11251G,11719A,11812G,13368A,14233G,14766T,14905A,15326G,15452A,15607G,15928A,16126C, 16167T,16261T,16294T,16296T,16304C,16519C
A11	K1a3a	73G,195C,263G,315.1C,497T,750G,1189C,1438G,1811G,2706G,3316A,3480G,4769G,6221C,7028T,7559G,8664G,8860G,9055A,9698C,10398G,10550G,11299C,11467G,11719A,12308G,12372A,13117G,14167T,14766T,14798C,15244G,15326G,16156A,16224C,16240C,16311C,16519C
A12	U6d	73G,263G,315.1C,750G,1438G,2706G,3348G,4336C,4769G,7028T,8860G,10397G,11467G,11719A,11947G,12308G,12372A12501A,12530G,14518G,14766T,15326G,16172C,16174T,16219G,16311C

A13	N1b1a5	73G,152C,263G,309.1C,315.1C,750G,1438G,1598A,1703T,1719A,2639T,2706G,3921A,4769G,4960T,4967C,5471A,7028T,8251A,8472T,8836G,8860G,9335T,10238C,11362G,11719A,12007A,12501A,12705T,12822G,14766T,15326G,16145A,16176G,16223T,16311C,16390A,16519C
A14	HV0b	72C,195C,198T,263G,315.1C,750G,1438G,2706G,4769G,7028T,8520G,8860G,10644A,15326G,16298C
A15	N1a1a1a2	73G,150T,152C,199C,204C,207A,263G,309.1C,315.1C,573.1C,573.2C,573.3C,573.4C,573.5C,573.6C,573.7C,669C,750G,1438G,1719A,2702A,2706G,3336C,4501T,4769G,5315G,7028T,8485A,8860G,8901G,10238C,10398G,11719A,11977T,12501A,12705T,13780G,14766T,15043A,15299C,15326G,16086C,16147A,16213A,16223T,16248T,16320T,16355T,16519C
A16	HV0b	72C,195C,198T,263G,315.1C,750G,1438G,2706G,4769G,7028T,8520G,8860G,10644A,15326G,16298C
A17	T2b11	73G,207A,263G,315.1C,709A,750G,930A,1438G,1888A,2706G,3398C,4216C,4580A,4769G,4902G,4917G,5147A,7028T,8557A,8697A,8860G,10463C,11251G,11719A,11812G,13368A,14233G,14766T,14905A,15326G,15452A,15607G,15928A,16126C,16167T,16261T,16294T,16296T,16304C,16519C
A18	T2b11	73G,207A,263G,315.1C,709A,750G,930A,1438G,1888A,2706G,3398C,4216C,4580A,4769G,4902G,4917G,5147A,7028T,8557A,8697A,8860G,9181G,10463C,11251G,11719A,11812G,13368A,14233G,14766T,14905A,15326G,15452A,15607G,15928A,16126C,16167T,16261T,16294T,16296T,16304C,16519C
A19	N1b1a2	73G,152C,204C,263G,309.1C,315.1C,750G,1438G,1598A,1703T,1719A,2639T,2706G,3921A,4769G,4904T,4960T,5471A,7028T,8251A,8472T,8836G,8860G,9335T,9947A,10238C,11362G,11719A,12501A,12705T,12822G,14766T,15326G,16145A,16176G,16223T,16390A,16519C

A20	U5a1a1	73G,263G,315.1C,750G,1438G,1700C,1819C,2706G,3197C,4769G,5495C,5806C,7028T,7670G,8860G,9477A,11467G, 11719A,12308G,12372A,13617C,14766T,14793G,15218G,15326G,15924G,16256T,16270T,16399G
A21	K1a3a	73G,195C,263G,315.1C,497T,750G,1189C,1438G,1811G,2706G,3316A,3480G,4769G,6221C,7028T,7559G,8664G,8860G, 9055A,9698C,10398G,10550G,11299C,11467G,11719A,12308G,12372A,13117G,14167T,14766T,14798C,15244G,15326G, 16156A,16224C,16240C,16311C,16519C
A22	HV0b	72C,195C,198T,263G,315.1C,750G,1438G,2706G,4769G,7028T,8520G,8860G,10644A,15326G,16298C
A23	H1	263G,315.1C,469T,750G,1438G,3010A,4769G,8860G,10007C,15326G,16092C,16172C,16519C
A24	T2e	73G,150T,263G,309.1C,315.1C,709A,750G,1438G,1888A,2308G,2706G,4216C,4769G,4917G,7028T,7133T,8697A,8860G, 10463C,11251G,11719A,11812G,13135A,13368A,14233G,14766T,14905A,15326G,15452A,15499T,15607G,15928A,16114T, 16126C,16153A,16192T,16294T,16519C

Region: Bragança
Longitude/Latitude: 41.48 / -6.45

Br01	H1	263G,315.1C,390G,750G,1438G,3010A,4769G,4790G,8860G,15326G,16519C
Br02	H1b	263G,315.1C,750G,1438G,3010A,4769G,8860G,15326G,16189C,16356C
Br03	X2b+226	73G,153G,195C,225A,226C,263G,315.1C,750G,813G,1438G,1719A,2706G,4044G,4769G,6221C,6371T,7028T,8393T,8860G,11719A,12705T,13708A,13966G,14470C,14766T,15326G,15927A,16183C,16189C,16223T,16278T,16519C
Br04	T2e1	41T,73G,150T,263G,309.1C,315.1C,709A,750G,1438G,1888A,2706G,4216C,4769G,4917G,7028T,8697A,8860G,9181G,10463C,11251G,11719A,11812G,13368A,14233G,14766T,14905A,15326G,15452A,15607G,15928A,16126C,16153A,16294T,16519C
Br05	H20a	249d,263G,292C,315.1C,750G,1438G,4769G,8860G,15326G,16218T,16328A,16362C
Br06	H3c	146C,263G,315.1C,573.1C,573.2C,750G,1438G,4769G,6776C,8860G,12957C,15326G,16519C

Br07	U5a1a1	73G,263G,315.1C,345T,750G,1438G,1700C,2706G,3197C,4769G,5468T,5495C,7028T,8860G,9477A,11467G,11719A,12308G,12372A,13617C,14020C,14766T,14793G,15218G,15326G,15924G,16126C,16256T,16270T,16399G
Br08	H3ag	263G,315.1C,750G,1438G,4769G,5498G,6293C,6776C,8860G,15326G,16519C
Br09	H3+152	152C,249d,263G,315.1C,750G,1438G,4769G,6776C,7965A,8860G,15326G,16519C
Br10	H1	263G,315.1C,750G,1438G,3010A,4769G,8860G,15326G,16519C
Br11	X2m'n	73G,146C,153G,195C,225A,226C,263G,309.1C,315.1C,750G,1438G,1719A,2706G,3918A,3987G,4769G,5201C,6221C,6371T,7028T,8860G,11719A,12705T,13966G,14470C,14766T,15314A,15326G,16189C,16223T,16255A,16278T,16519C
	Region: Carção	
	Longitude/Latitude: 41.35 / -6.35	
Ca1	N1a1a1a2	73G,150T,152C,199C,204C,207A,263G,309.1C,315.1C,573.1C,573.2C,573.3C,573.4C,573.5C,573.6C,573.7C,573.8C,669C,750G,1438G,1719A,2702A,2706G,3336C,4501T,4769G,5315G,7028T,8485A,8860G,8901G,10238C,10398G,11719A,11977T,12501A,12705T,13780G,14766T,15043A,15299C,15326G,16086C,16147A,16223T,16248T,16320T,16355T,16519C

Ca2	U5b1b1 +152	73G,150T,152C,263G,315.1C,523d,524d,750G,1438G,1822C,2706G,3197C,4769G,5656G,5788C,7028T,7385G,7768G,8308G, 8860G,9477A,10301G,10927C,11467G,11719A,12308G,12372A,12618A,13617C,14182C,14766T,15326G,15533G,16074G, 16189C,16192T,16270T
Ca3	H23	214G,263G,315.1C,750G,1438G,4769G,5999C,6032A,8839A,8860G,10101C,10211T,14569A,15326G,16519C
Ca4	HV0b	72C,195C,198T,263G,315.1C,750G,1438G,2706G,4769G,7028T,8520G,8860G,10644A,15326G,16298C
Ca5	U2e1a1	73G,152C,217C,263G,315.1C,340T,508G,524.1A,524.2C,750G,1438G,1811G,2706G,3116T,3720G,4769G,5390G,5426C,6045T, 6152C,7028T,8014G,8860G,10876G,11197T,11467G,11719A,11732C,12308G,12372A,13020C,13708A,13734C,14766T,15326G, 15907G,16051G,16129C,16183C,16189C,16362C,16519C
Ca6	HV0b	72C,195C,198T,263G,309.1C,315.1C,750G,1438G,2706G,4769G,7028T,8520G,8860G,10644A,11914A,15326G,16298C
Ca7	U2e1a1	73G,152C,217C,263G,315.1C,340T,508G,524.1A,524.2C,750G,1438G,1811G,2706G,3116T,3720G,4769G,5390G,5426C,6045T, 6152C,7028T,8014G,8860G,10876G,11197T,11467G,11719A,11732C,12308G,12372A,13020C,13708A,13734C,14766T,15326G, 15907G,16051G,16129C,16183C,16189C,16193.1C,16362C,16519C
Ca8	U2e1a1	73G,152C,217C,263G,315.1C,340T,508G,524.1A,524.2C,750G,1438G,1811G,2706G,3116T,3720G,4769G,5390G,5426C,6045T, 6152C,7028T,8014G,8860G,10876G,11197T,11467G,11719A,11732C,12308G,12372A,13020C,13708A,13734C,14766T,15326G, 15907G,16051G,16129C,16183C,16189C,16362C,16519C

Region: Mogadouro		
Longitude/Latitude: 41.20 / -6.42		
Mo1	H42	263G,309.1C,315.1C,750G,1438G,4769G,8860G,9758C,15326G,15625T,16519C
Region: Vilarinho dos Galegos		
Longitude/Latitude: 41.16 / -6.37		
Vi01	U8a1a1a	73G,263G,282C,309.1C,315.1C,750G,1438G,1811G,2706G,3738T,4129G,4769G,5240G,6392C,6455T,7028T,7055G,8860G,9365T9698C,10733T,11467G,11719A,12135A,12308G,12372A,13145A,14766T,15326G,16209C,16342C
Vi02	H1	263G,315.1C,750G,1438G,3010A,4769G,8860G,10007C,12811C,15326G,16172C,16234T,16235G,16399G,16519C
Vi03	H	93G,200G,263G,315.1C,750G,4769G,1438G,3254T,8860G,15326G,16519C
Vi04	V+150	44.1C,72C,150T,263G,309.1C,315.1C,750G,1438G,2706G,4580A,4769G,7028T,8860G,13392C,15326G,15904T,15928A,16298C

V.i05	H	93G,200G,263G,315.1C,750G,1438G,3254T,4769G,8860G,15326G,16519C
V.i06	H13a1	153G,263G,315.1C,523d,524d,750G,1438G,2259T,4231G,4745G,4769G,7903G,8860G,14872T,15326G,16291T,16519C
V.i07	H1	263G,315.1C,750G,1438G,3010A,4769G,8860G,10007C,12811C,15326G,16172C,16234T,16235G,16399G,16519C
V.i08	U2e1a1	73G,152C,217C,263G,315.1C,340T,508G,524.1A,524.2C,750G,1438G,1811G,2706G,3116T,3720G,4769G,5390G,5426C,6045T,6152C,7028T,8014G,8860G,10876G,11197T,11467G,11719A,11732C,12308G,12372A,13020C,13708A,13734C,14766T,15326G,15907G,16051G,16129C,16183C,16189C,16193.1C,16362C,16519C
V.i09	J1b2	73G,263G,295T,315.1C,462T,489C,750G,1438G,1733T,2706G,3010A,4216C,4769G,6719C,7028T,8269A,8860G,10398G,11251G,11719A,12612G,13708A,14766T,14927G,15326G,15452A,16069T,16126C,16145A,16222T,16261T
V.i10	H1	263G,315.1C,750G,1438G,3010A,4769G,8860G,10007C,12811C,15326G,16172C,16235G,16399G,16519C
V.i11	U5b1c	73G,150T,194T,263G,309.1C,315.1C,750G,1438G,2706G,3197C,4769G,5656G,6216C,7028T,7080C,7768G,8014G,8860G,9477A,10398G,10586A,11467G,11719A,12308G,12372A,13617C,14182C,14766T,15191C,15326G,16172C,16189C,16192T,16270T,16311C,16519C

Vi12	H1	263G,315.1C,750G,1438G,3010A,4769G,8860G,15326G,15799G,16519C
Vi13	H1	263G,315.1C,750G,1438G,3010A,4769G,8860G,10007C,12811C,15326G,16234T,16235G,16399G,16519C

Table S2. Haplogroup frequencies of Jews from Bragança together with the Portuguese host population and other geographical neighbours. Haplogroup definition based on Phylotree, mtDNA tree Build 16, considering variations of the control region (16024 – 16400 bp for HVRI, and 073 – 340 bp for HVRII).

Haplogroup	Population N (%)			
	Portuguese Jews	Portugal	Miranda	NW Spain
B6a1a	-	2 (0.4)	-	-
H1+16189	-	7 (1.3)	-	1 (0.5)
H10a1b	-	1 (0.2)	-	0
H10e	-	1 (0.2)	-	-
H11a	-	1 (0.2)	-	-
H13a1a2a	-	1 (0.2)	-	-
H13a1b	1 (1.8)	-	-	1 (0.5)
H13a1d	-	-	-	1 (0.5)
H13a2c	-	-	-	1 (0.5)
H14b1	-	2 (0.4)	-	-
H15a1b	-	1 (0.2)	-	-
H17c	-	1 (0.2)	-	-
H1a	-	-	1 (0.8)	-
H1a1	-	2 (0.4)	-	1 (0.5)
H1a3	-	1 (0.2)	-	-
H1ag1a	-	1 (0.2)	-	-
H1aj1	-	4 (0.7)	-	-
H1an2	-	1 (0.2)	-	-
H1ap1	-	2 (0.4)	-	-
H1b	1 (1.8)	-	-	2 (0.9)
H1b1+16362	-	2 (0.4)	-	1 (0.5)
H1b3	-	2 (0.4)	-	-
H1ba	-	2 (0.4)	-	-
H1bv1	-	3 (0.5)	1 (0.8)	-
H1bf1	-	1 (0.2)	-	-
H1bo	-	-	1 (0.8)	-
H1c+152	-	22 (4)	7 (5.8)	4 (1.9)
H1c1	-	2 (0.4)	1 (0.8)	-
H1c3	-	5 (0.9)	-	-
H1c3a	-	1 (0.2)	-	-
H1c4b	-	1 (0.2)	-	-
H1c5a	1 (1.8)	-	-	-
H1e+16129	-	5 (0.9)	-	6 (2.8)
H1e1a1	2 (3.5)	4 (0.7)	-	-
H1e1a4	-	1 (0.2)	-	3 (1.4)
H1e1a6	-	6 (1.1)	3 (2.5)	2 (0.9)
H1e2c	-	3 (0.5)	3 (2.5)	4 (1.9)

Table S2 – (Continued)

H1e4a	-	1 (0.2)	1 (0.8)	-
H1e5	-	-	-	1 (0.5)
H1j2a	-	1 (0.2)	-	-
H1j8	-	1 (0.2)	-	-
H1k	-	1 (0.2)	-	-
H1m1	1 (1.8)	3 (0.5)	3 (2.5)	4 (1.9)
H1n+146+195	-	2 (0.4)	-	-
H1r1	-	-	-	1 (0.5)
H1x	-	-	4 (3.3)	1 (0.5)
H20a	1 (1.8)	-	1 (0.8)	1 (0.5)
H24	-	2 (0.4)	-	-
H26b	-	3 (0.5)	-	1 (0.5)
H27	-	2 (0.4)	1 (0.8)	-
H2a1	-	2 (0.4)	-	-
H2a2a	9 (15.8)	71 (12.9)	15 (12.4)	34 (15.9)
H2a2a+(16235)	3 (5.3)	-	1 (0.8)	-
H2a2b	-	4 (0.7)	3 (2.5)	-
H2a3	-	3 (0.5)	-	2 (0.9)
H32	-	2 (0.4)	-	-
H39	-	1 (0.2)	-	1 (0.5)
H39a	-	1 (0.2)	-	-
H3an	-	1 (0.2)	-	-
H3c2	-	1 (0.2)	-	1 (0.5)
H3s	-	1 (0.2)	-	-
H3v+16093	-	4 (0.7)	3 (2.5)	1 (0.5)
H3x	-	-	3 (2.5)	-
H3z	-	1 (0.2)	-	2 (0.9)
H41a	-	1 (0.2)	-	-
H4a1a+195	-	3 (0.5)	3 (2.5)	2 (0.9)
H5	-	11 (2)	4 (3.3)	-
H5a1j	-	3 (0.5)	-	2 (0.9)
H5b1	1 (1.8)	-	-	-
H6	-	5 (0.9)	3 (2.5)	-
H66a	-	3 (0.5)	-	1 (0.5)
H6a1a1a	1 (1.8)	-	1 (0.8)	-
H6a1a7	-	1 (0.2)	2 (1.7)	-
H7a1a	-	1 (0.2)	-	-
H7c1	-	1 (0.2)	-	-
H7c4	-	2 (0.4)	-	-
H7h	-	1 (0.2)	-	-
H8	-	-	-	1 (0.5)
H82	-	-	-	1 (0.5)
H96	-	2 (0.4)	-	-
HV0	-	17 (3.1)	6 (5)	8 (3.7)

Table S2 – (Continued)

HV0+195	-	8 (1.5)	-	3 (1.4)
HV0b	7 (12.3)	-	-	-
HV4a	-	1 (0.2)	2 (1.7)	-
HV9c	-	-	-	1 (0.5)
I	-	3 (0.5)	-	-
I1a1	-	6 (1.1)	-	-
I1d	-	2 (0.4)	-	-
I4a1	-	-	-	3 (1.4)
J	-	2 (0.4)	-	3 (1.4)
J1b	1 (1.8)	2 (0.4)	5 (4.1)	1 (0.5)
J1b1a1	-	3 (0.5)	2 (1.7)	2 (0.9)
J1b3a	-	1 (0.2)	-	-
J1c	-	5 (0.9)	1 (0.8)	3 (1.4)
J1c+16261	-	2 (0.4)	-	-
J1c+16261+189	-	1 (0.2)	-	-
J1c15b	-	1 (0.2)	-	-
J1c1e	-	1 (0.2)	-	-
J1c2	-	3 (0.5)	2 (1.7)	4 (1.9)
J1c2c1	-	2 (0.4)	-	1 (0.5)
J1c2e	-	-	-	2 (0.9)
J1c2e1	-	-	-	1 (0.5)
J1c2h	-	2 (0.4)	-	-
J1c2m	-	-	-	1 (0.5)
J1c3f	-	1 (0.2)	-	-
J1c3j	-	2 (0.4)	1 (0.8)	-
J1c8a	-	3 (0.5)	-	-
J2a1a1	-	-	-	2 (0.9)
J2a2	-	1 (0.2)	-	-
J2a2b1	-	-	1 (0.8)	-
J2b	-	1 (0.2)	-	-
J2b1a	-	1 (0.2)	-	6 (2.8)
J2b1c1	-	3 (0.5)	-	-
JT	-	1 (0.2)	-	-
K	2 (3.5)	10 (1.8)	3 (2.5)	6 (2.8)
K1a	-	4 (0.7)	-	-
K1a+150	-	1 (0.2)	-	1 (0.5)
K1a13	-	1 (0.2)	-	-
K1a1b2a	-	1 (0.2)	-	-
K1a2a1	-	1 (0.2)	-	1 (0.5)
K1a4a1a+195	-	4 (0.7)	-	-
K1a4c	-	2 (0.4)	-	-
K1b1a	-	-	-	1 (0.5)
K2b1a1a	-	-	-	3 (1.4)
K1b2	-	3 (0.5)	-	-

Table S2 – (Continued)

K1c	-	3 (0.5)	-	-
K2	-	1 (0.2)	-	-
L1b	-	4 (0.7)	-	1 (0.5)
L1b1a+189	-	1 (0.2)	-	-
L1b1a1'4	-	1 (0.2)	-	-
L1b1a12	-	1 (0.2)	1 (0.8)	3 (1.4)
L1b1a2a	-	2 (0.4)	-	-
L1c1'2'4'5'6	-	1 (0.2)	-	-
L1c2b1c	-	1 (0.2)	-	-
L2a1+143	-	2 (0.4)	-	-
L2a1a3	-	1 (0.2)	1 (0.8)	-
L2a1b+143	-	1 (0.2)	-	-
L2a1b3	-	1 (0.2)	-	-
L2a1c+16129	-	-	4 (3.3)	-
L2b	-	1 (0.2)	-	2 (0.9)
L3b1a+@16124	-	1 (0.2)	-	4 (1.9)
L3d1b3a	-	1 (0.2)	-	-
L3e'i'k'x	-	4 (0.7)	-	-
L3e1a1	-	1 (0.2)	-	-
L3e2b	-	1 (0.2)	-	-
L3e5	-	1 (0.2)	-	-
L3f1b	-	2 (0.4)	-	-
L3f1b4a	-	1 (0.2)	-	-
L3f1b4c	-	1 (0.2)	-	-
L3x2b	-	1 (0.2)	-	-
M1	-	1 (0.2)	-	-
M1a1	-	1 (0.2)	-	-
M1a3	-	1 (0.2)	-	1 (0.5)
M1b1a	-	1 (0.2)	-	-
M5a1	-	-	1 (0.8)	-
M8	-	1 (0.2)	-	-
N1a	-	-	-	1 (0.5)
N1a1a1a2	2 (3.5)	1 (0.2)	-	-
N1a1b	-	1 (0.2)	-	-
N1b1	2 (3.5)	-	5 (4.1)	2 (0.9)
N1b1a+195	-	1 (0.2)	-	-
N1b1b	-	1 (0.2)	-	-
R0+16189	-	3 (0.5)	-	-
R0a	-	2 (0.4)	-	1 (0.5)
R1	-	1 (0.2)	-	-
R30a1b	-	1 (0.2)	-	-
R5	-	2 (0.4)	-	-
R5a	-	1 (0.2)	-	-
R8	-	1 (0.2)	-	-

Table S2 – (Continued)

R8a1a1b	-	1 (0.2)	-	-
R9c1	-	1 (0.2)	-	-
T	-	4 (0.7)	-	5 (2.3)
T1	-	3 (0.5)	-	-
T1a	-	1 (0.2)	1 (0.8)	2 (0.9)
T1a+152	-	1 (0.2)	-	-
T1a1'3	-	3 (0.5)	-	-
T1a2a	-	2 (0.4)	-	-
T1a9	-	3 (0.5)	-	-
T2	-	2 (0.4)	-	-
T2a1b	-	1 (0.2)	-	-
T2b	-	3 (0.5)	2 (1.7)	1 (0.5)
T2b+150	-	1 (0.2)	-	-
T2b11	4 (7)	-	-	1 (0.5)
T2b13a	-	5 (0.9)	-	-
T2b23	-	1 (0.2)	-	-
T2b24	-	1 (0.2)	-	-
T2b3+151	-	10 (1.8)	-	-
T2c1d+152	-	1 (0.2)	-	-
T2d1	-	1 (0.2)	-	-
T2e	1 (1.8)	3 (0.5)	-	-
T2e1a1a1	2 (3.5)	1 (0.2)	1 (0.8)	-
T2f2	-	1 (0.2)	-	-
T2f3	-	-	1 (0.8)	-
T2j	-	2 (0.4)	-	-
U2	-	-	-	1 (0.5)
U2d3	-	-	-	1 (0.5)
U2e1	4 (7)	4 (0.7)	2 (1.7)	-
U2e1b1	-	1 (0.2)	-	-
U2e1h	-	-	-	3 (1.4)
U2e2	-	2 (0.4)	-	-
U3	-	2 (0.4)	-	-
U3a	-	1 (0.2)	-	-
U3a1c	-	2 (0.4)	-	-
U4a1	-	1 (0.2)	-	-
U4a2	-	2 (0.4)	-	-
U4c1	-	5 (0.9)	-	-
U4d2	-	1 (0.2)	-	-
U5a	-	3 (0.5)	-	2 (0.9)
U5a1	-	-	-	2 (0.9)
U5a1+@16192	3 (5.3)	1 (0.2)	-	1 (0.5)
U5a1a1+152	1 (1.8)	-	-	-
U5a1b+16362	-	-	-	1 (0.5)
U5a1a1d	-	1 (0.2)	-	1 (0.5)

Table S2 – (Continued)

U5a1b1	-	2 (0.4)	1 (0.8)	-
U5a1d2a	-	-	-	1 (0.5)
U5a1g1	-	-	-	1 (0.5)
U5a1h	-	-	-	1 (0.5)
U5a2+16362	-	2 (0.4)	-	-
U5a2a	-	1 (0.2)	-	2 (0.9)
U5a2c3	-	4 (0.7)	-	-
U5b	1 (1.8)	-	-	4 (1.9)
U5b1+16189	-	3 (0.5)	-	2 (0.9)
U5b1b1+@16192	-	4 (0.7)	2 (1.7)	1 (0.5)
U5b1b1+152	1 (1.8)	5 (0.9)	-	-
U5b1c	-	1 (0.2)	-	-
U5b1c1	-	1 (0.2)	-	-
U5b1d1a	-	-	2 (1.7)	-
U5b1d1b	-	-	-	1 (0.5)
U5b1d2	-	-	1 (0.8)	-
U5b1e	-	1 (0.2)	-	-
U5b1f1a	-	-	-	1 (0.5)
U5b1g	-	1 (0.2)	-	-
U5b1i	-	3 (0.5)	-	1 (0.5)
U5b2a1a+16311	-	1 (0.2)	-	3 (1.4)
U5b2a1a2	-	1 (0.2)	-	1 (0.5)
U5b2a5	-	1 (0.2)	1 (0.8)	3 (1.4)
U5b2b1a1	-	1 (0.2)	-	2 (0.9)
U5b2b3	-	-	-	2 (0.9)
U5b2b3a	-	2 (0.4)	-	-
U5b3	-	-	-	1 (0.5)
U5b3a	-	1 (0.2)	-	-
U6	-	-	-	1 (0.5)
U6a'b'd+16311	1 (1.8)	2 (0.4)	-	-
U6a1a	-	3 (0.5)	-	-
U6a1a1	-	3 (0.5)	-	-
U6a1b1a	-	3 (0.5)	-	-
U6a1b1b	-	1 (0.2)	-	-
U6a1b2	-	1 (0.2)	-	-
U6a3+185	-	1 (0.2)	-	-
U6a7a	-	3 (0.5)	-	-
U6d1	-	1 (0.2)	-	-
U6d3	-	1 (0.2)	2 (1.7)	-
U7	-	1 (0.2)	-	-
U8a	1 (1.8)	1 (0.2)	-	-
U8a1a	-	-	-	2 (0.9)
V+150	1 (1.8)	-	4 (3.3)	-
V16	-	1 (0.2)	-	5 (2.3)

Table S2 – (Continued)

V3a	-	1 (0.2)	-	2 (0.9)
V7a	-	-	-	1 (0.5)
V9a	-	1 (0.2)	-	-
W	-	2 (0.4)	-	-
W+194	-	-	1 (0.8)	2 (0.9)
W1+119	-	3 (0.5)	-	-
W1g	-	1 (0.2)	-	-
W4	-	1 (0.2)	-	-
W5a	-	1 (0.2)	-	-
W6	-	1 (0.2)	-	-
X1	-	1 (0.2)	-	-
X1'3	-	1 (0.2)	-	-
X2+225	-	1 (0.2)	-	-
X2+225+@153	-	2 (0.4)	-	-
X2b+226	2 (3.5)	8 (1.5)	-	1 (0.5)
X2c	-	2 (0.4)	-	-
X2d	-	1 (0.2)	-	-
Total	**57**	**549**	**121**	**214**

Supplemental figures and legends

Figure S1. Control Region MJ network of haplogroup HV0.

Haplogroup HV0 defining motif is shown in brackets and corresponding haplotype is marked with a red circle. Numbers on the branches refer to HVRI and HVRII (16024-16365 bp and 72-300 bp, respectively) substitutions referred to the revised Cambridge Reference Sequence (rCRS), circle sizes are proportional to haplotype frequencies.

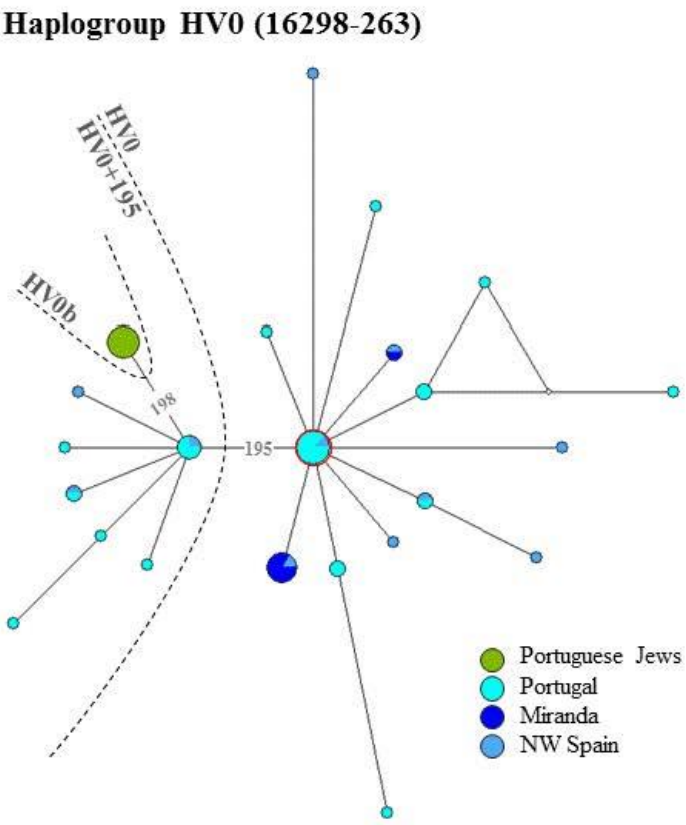


Figure S2. Detailed maximum parsimony trees of 7 novel complete Portuguese Jews mtDNA sequences belonging to haplogroup HV0b.

Tree includes samples from public databases (NCBI-GeneBank, empop.org and mtDNACommunity.org). In green, samples from the present work; in orange, a Jewish Sephardi sample; in blue, other samples. Variants relative to the rCRS are shown on the branches; they are transitions unless a base is explicitly indicated. Suffixes indicate: transversions (to A, G, C, or T), indels (.1, d), gene locus (,r, rRNA; ,t, tRNA), synonymous or non-synonymous changes (s or ns), non-coding sites outside the control region (nc) and back variants to an ancestral state (!). Recurrent variants within the phylogeny are underlined. Neither the variation in number of cytosines around nps 309 and 16193, nor heteroplasmic positions were considered in the tree.

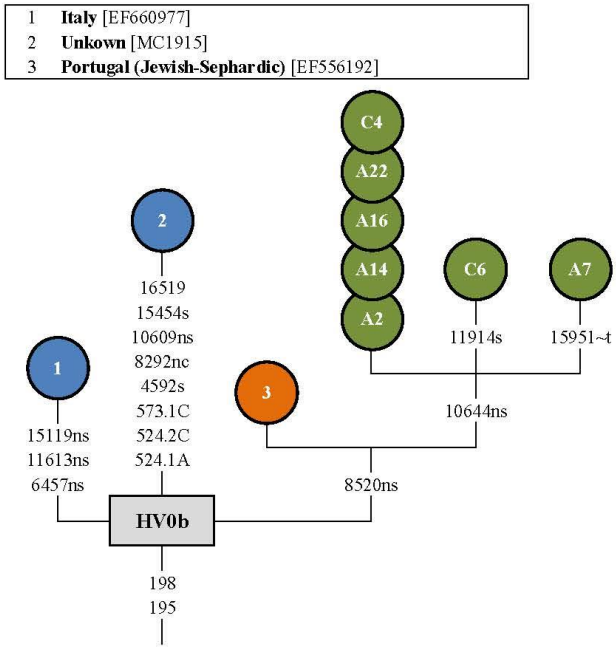


Figure S3. Control Region MJ network of haplogroups N1a and N1b1.

Haplogroups N1a and N1b1 defining motifs are shown in brackets and corresponding haplotypes are marked with red circles. Numbers on the branches refer to HVRI and HVRII (16024-16365 bp and 72-300 bp, respectively) substitutions referred to the revised Cambridge Reference Sequence (rCRS), circle sizes are proportional to haplotype frequencies.

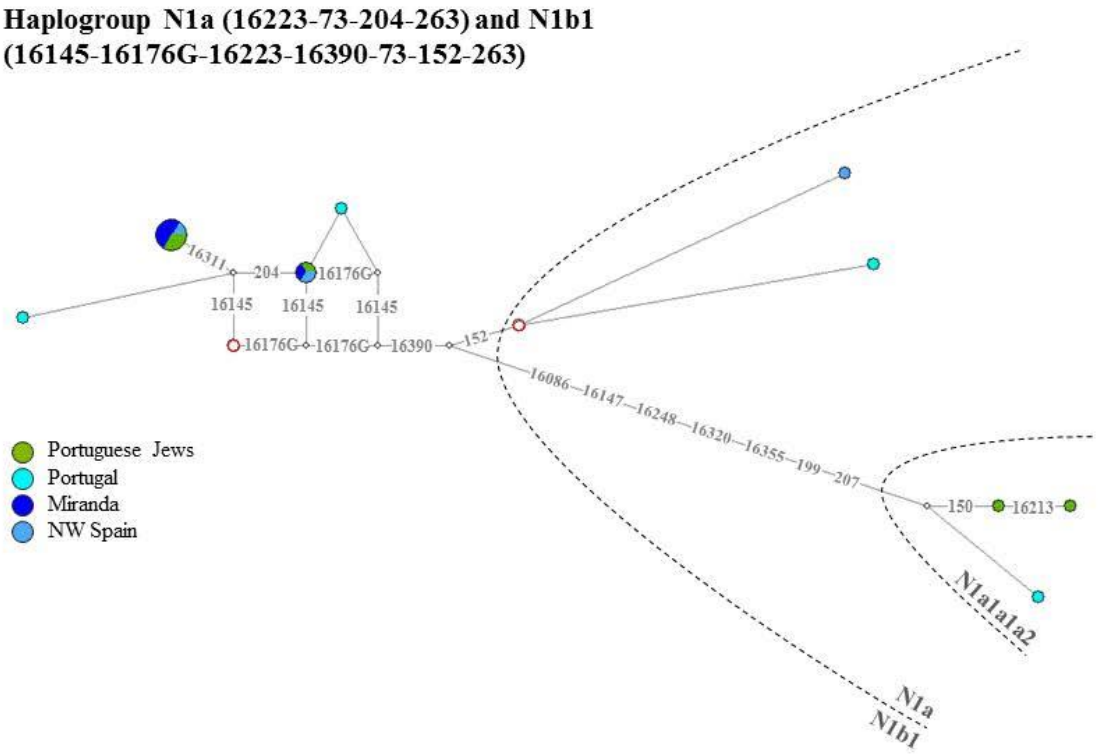


Figure S4. Detailed maximum parsimony trees of 4 novel complete Portuguese Jews mtDNA sequences belonging to haplogroup N1 (N1a1a1a2 and N1b1a).

Trees include samples from public databases (NCBI-GeneBank, empop.org and mtDNACommunity.org). In green, samples from the present work; in blue, other samples. Variants relative to the rCRS are shown on the branches; they are transitions unless a base is explicitly indicated. Suffixes indicate: transversions (to A, G, C, or T), indels (.1, d), gene locus (,r, rRNA; ,t, tRNA), synonymous or non-synonymous changes (s or ns), non-coding sites outside the control region (nc) and back variants to an ancestral state (!). Recurrent variants within the phylogeny are underlined. Neither the variation in number of cytosines around nps 309 and 16193, nor heteroplasmic positions were considered in the trees.

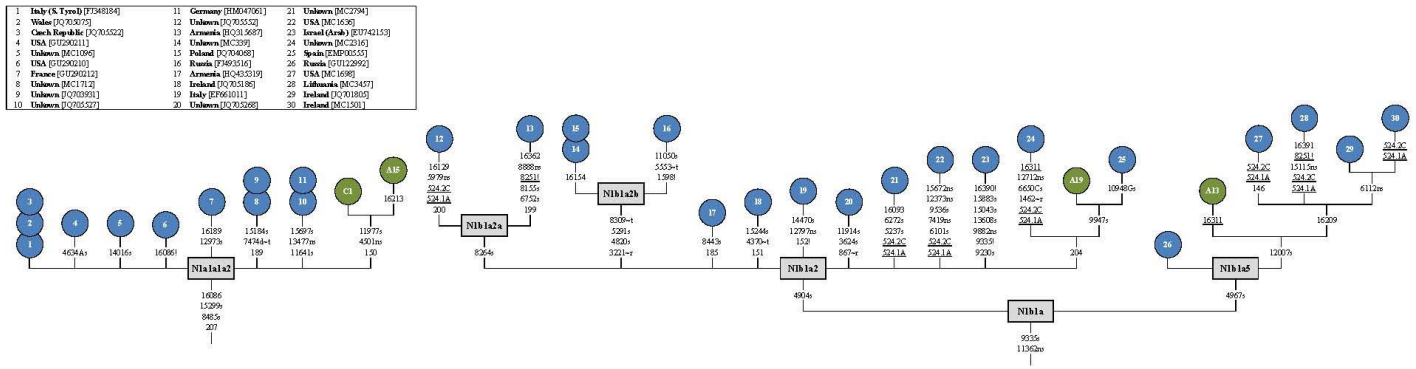


Figure S5. Control Region MJ network of haplogroup T2b.

Haplogroup T2b defining motif is shown in brackets and corresponding haplotype is marked with a red circle. Numbers on the branches refer to HVRI and HVRII (16024-16365 bp and 72-300 bp, respectively) substitutions referred to the revised Cambridge Reference Sequence (rCRS), circle sizes are proportional to haplotype frequencies.

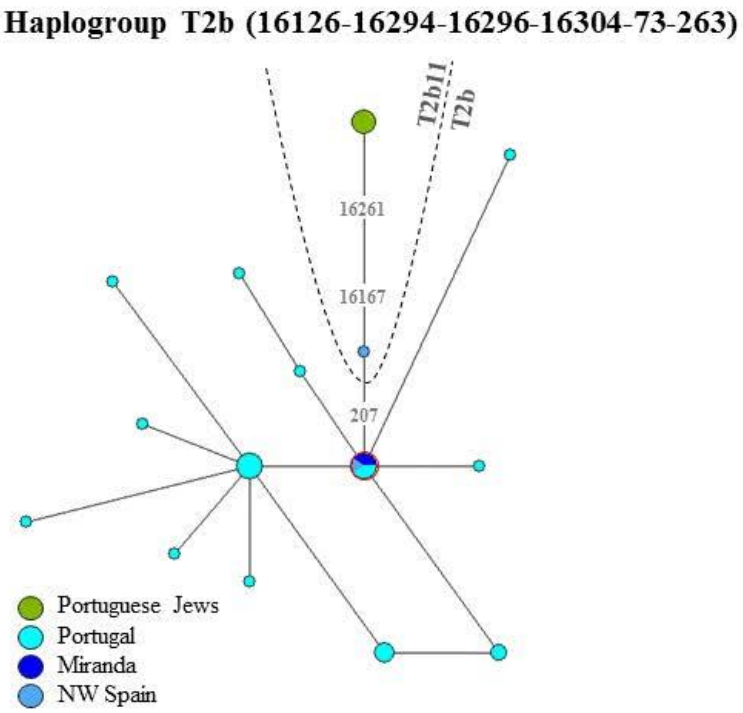


Figure S6. Detailed maximum parsimony trees of 7 novel complete Portuguese Jews mtDNA sequences belonging to haplogroup T2 (T2b11 and T2e1).

Trees include samples from public databases (NCBI-GeneBank, empop.org and mtDNACommunity.org). In green, samples from the present work; in orange, Jewish samples, Sephardim or Ashkenazim; in blue, other samples. Variants relative to the rCRS are shown on the branches; they are transitions unless a base is explicitly indicated. Suffixes indicate: transversions (to A, G, C, or T), indels (.1, d), gene locus (,r, rRNA; ,t, tRNA), synonymous or non-synonymous changes (s or ns), non-coding sites outside the control region (nc) and back variants to an ancestral state (!). Recurrent variants within the phylogeny are underlined. Neither the variation in number of cytosines around nps 309 and 16193, nor heteroplasmic positions were considered in the trees.

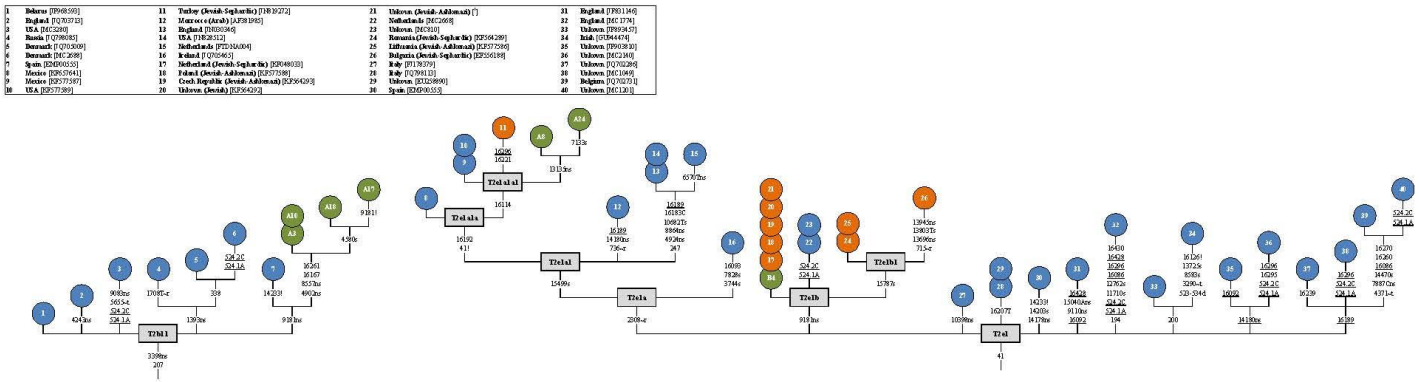


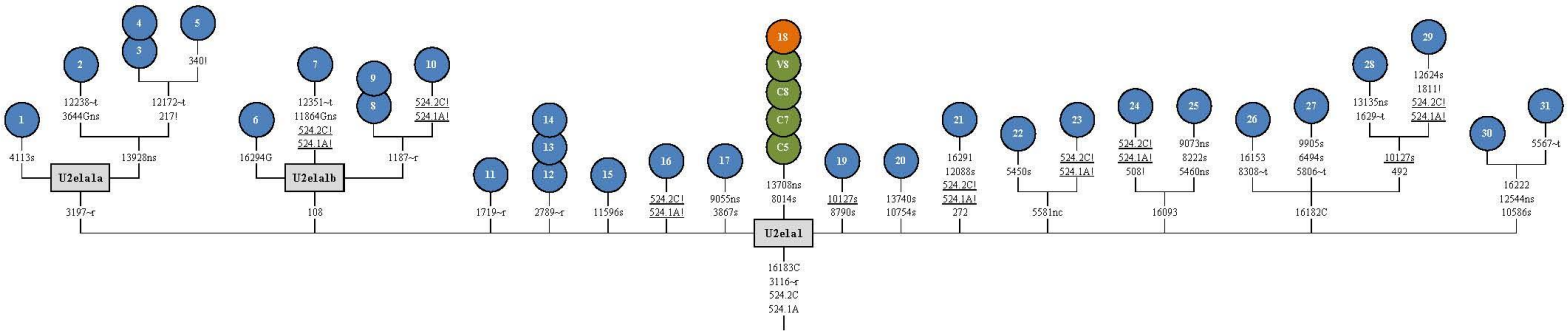
Figure S7. Control Region MJ network of haplogroup U2e1.

Haplogroup U2e1 defining motif is shown in brackets and corresponding haplotype is marked with a red circle. Numbers on the branches refer to HVRI and HVRII (16024-16365 bp and 72-300 bp, respectively) substitutions referred to the revised Cambridge Reference Sequence (rCRS), circle sizes are proportional to haplotype frequencies.

Figure S8. Detailed maximum parsimony trees of 4 novel complete Portuguese Jews mtDNA sequences belonging to haplogroup U2e1a1.

Tree including samples from public databases (NCBI-GeneBank, empop.org and mtDNACommunity.org). In green, samples from the present work; in orange, a Jewish Ashkenazi sample; in blue, other samples. Variants relative to the rCRS are shown on the branches; they are transitions unless a base is explicitly indicated. Suffixes indicate: transversions (to A, G, C, or T), indels (.1, d), gene locus (,r, rRNA; ,t, tRNA), synonymous or non-synonymous changes (s or ns), non-coding sites outside the control region (nc) and back variants to an ancestral state (!). Recurrent variants within the phylogeny are underlined. Neither the variation in number of cytosines around nps 309 and 16193, nor heteroplasmic positions were considered in the tree.

1	**Unkown** [JQ705936]	11	**Unkown** [JQ702398]	21	**Unkown** [JQ703957]	31	**Unkown** [JQ705602]
2	**Finland** [AY339545]	12	**Ireland** [MC444]	22	**France** [JQ703530]		
3	**USA** [FJ490554]	13	**Ireland** [MC1351]	23	**Portugal** [EF177434]		
4	**Unkown** [JQ705117]	14	**USA** [HQ218053]	24	**Unkown** [JQ705989]		
5	**Finland** [AY339546]	15	**Unkown** [JQ703589]	25	**Unkown** [MC441]		
6	**Unkown** [JQ705861]	16	**Spain** [AY882382]	26	**India** [AY714040]		
7	**Unkown** [JQ703203]	17	**Unkown** [JQ705559]	27	**Unkown** [JQ705711]		
8	**Germany** [JQ704999]	18	**Moldova (Jewish-Ashkenazi)** [JQ702106]	28	**Italy (S. Tyrol)** [FJ348194]		
9	**Unkown** [MC644]	19	**Unkown** [JQ704817]	29	**Italy (S. Tyrol)** [FJ348188]		
10	**USA** [HQ424846]	20	**Unkown** [JQ704061]	30	**Ireland** [MC696]		



Letter:

Nogueiro, I., Teixeira, J., Amorim, A., Gusmão, L., & Alvarez, L. (2014). Reply to letter from Felice L. Bedford and Doron Yacobi. *European Journal of Human Genetics*.

European Journal of Human Genetics (2014), 1–2

www.nature.com/ejhg

LETTER

Reply to letter from Felice L. Bedford and Doron Yacobi

European Journal of Human Genetics advance online publication, 5 November 2014; doi:10.1038/ejhg.2014.232

The authors of Nogueiro *et al*[1] do thank Bedford and Yacobi comments on our paper providing the opportunity to clarify some issues that may have been put forward in an insufficiently clear or poorly phrased manner.

We must begin by restating the framework in which the paper was worked out; considering that the Iberian Peninsula constitutes the original geographic/historical source of Sephardic populations, the main goal was to characterize complete mitogenomes from self-designated Jews from north-eastern Portugal. The pertinence of the study lays on that the majority of previous studies (Belmonte[2] and Mallorca[3] excluded) were based in descents from exiled Sephardic communities with a supposed Iberian origin rather than those who stayed in Iberia and constitute the remains of the original Sephardic population.

Accordingly, all previous works reporting maternal Sephardic lineages were addressed including those published by Bedford[4] and Bedford *et al*.[5]

The clades of all haplotypes in Nogueiro *et al*[1] work were assigned according to PhyloTree built 16.[6] This nomenclature was adopted during the revision process of the manuscript (see revision track record of Nogueiro *et al*[1] work), since the last version of the PhyloTree was launched on the 19 February 2014. At no moment, we intended to credit to us the authorship of the definition of any new branch of the mitochondrial DNA (mtDNA) phylogenetic tree, as clearly stated in the text.

The purpose of the phylogenetic tree of global human mtDNA variation and haplogroup nomenclature (www.phylotree.org) is to provide, as stated by the curators, a framework to the scientific community with update information from novel mitogenomes sequences.[6] This tool has greatly improved the way researchers can transmit, compare and contribute with their results to the global mtDNA phylogeny knowledge. In this sense, as far as we understand, when a specific PhyloTree built is cited, it is credited to all contributors, as references of mitogenomes upon which a specific branch is based are always quoted.

In consequence, we do not understand the purpose of the Bedford and Yacobi[7] letter concerning the credit of authorship attribution of the two Jewish clades. We do think to have followed the standard procedures in this matter but we apologize if, by any means that we have not intended, our text can be interpreted as implying an inappropriate authorship attribution.

Regarding the first specific branch addressed in Bedford and Yacobi,[7] T2e1b (and T2e1b1), as referred in Nogueiro *et al*,[1] was newly included in the PhyloTree build 16 (van Oven and Kayser[6]) based on three quoted mitogenomes (GenBank accession numbers: KF048033, KF577586 and EF556188). In fact, Bedford *et al*[5] submitted two of these mitogenomes, the remaining one being contributed by Behar *et al*.[2] Although the claims contained in Bedford and Yacobi[7] could start a more transversal debate, we consider that in this specific issue no further acknowledgment should be expected as branches of the mitochondrial phylogenetic tree are not sponsored/personal domains. Otherwise, the detailed history of any (sub)haplogroup definition would have to be at least summarised, which in fact, for the one under discussion would require a lengthy and tedious description beginning with Torroni *et al*[8] and including all the works having contributed for the definition of the branch.

Concerning the complete mitogenomes used to construct the most parsimonious tree of the analysed clades in Nogueiro *et al*[1] (Supplementary Figures 2), all the information regarding accession numbers (GenBank and EMPOP), sample IDs (mDNA community) and/or direct citations is included. For example, following the corresponding links of any accession number, the information publicly available at GenBank includes: Authors, Title, Journal, Location and Ethnicity, among other details. Whenever ambiguous information was found, we have contacted by email the respective authors for clarification as it happened indeed with Felice L. Belford among others. This added information was duly mentioned in the Acknowledgments section. Furthermore, in no part of the work of Nogueiro *et al*,[1] the ethnicity of the samples used for comparisons is credited to us. Again, Bedford and Yacobi[7] comments concerning this topic could lead to an interesting discussion focused on the presence of shared female lineages between Bedford[4] and Bedford *et al*[5] previous works and the Portuguese Jews.

Considering now T2e1a1a, as for T2e1b, a new branch in the PhyloTree build 16 (van Oven and Kayser[6]) was included. This branch was defined based on a mitogenome (GenBank accession number: KF657641) from Bedford *et al*[5] as can be observed in PhyloTree build 16 search. This is a sample from Mexico: Nuevo Leon and together with three other samples from Bedford *et al*[4,5] (KF577587, KF577589 and JN819272) establish a clade where the Portuguese Sephardic Jews are included (T2e1a1a1) in Nogueiro *et al*[1] (Supplementary Figure 6). With the exception of JN819272, a Sephardic Jew from Salonica, the other two samples are from Mexico, Tamaulipas and from Texas, USA and could have an eventual Iberian/Sephardic origin. The fact that we have observed two Jewish Sephardic sequences in the NE Portugal sharing the same motif described by Bedford *et al*,[4,5] plus two further distinct variants, m.13135G4A and m.7133C4T not previously described, can help to clarify if the putative Iberian origin of the Mexican/USA samples is indeed owing to Sephardic ancestry.

It is remarkable that T2e1b and U2e1a clades include mitogenomes from both Sephardic and Ashkenazi origins. We do recognise that Bedford *et al*[5] have suggested the two possible scenarios for the T2e1b clade that we also consider, either the defining variants for each branch could have arisen before the separation between the two Jewish groups, or there may have been recent admixture between them. However, in Nogueiro *et al*[1] it is stated that: 'More complete sampling and complete sequences will contribute to the clarification of which one is more likely. In any case, it must be said that although

not frequent, marriages between the two communities occurred (especially) in the sixteenth and seventeenth centuries, namely among the elite sugar traders, with the descendants assimilated into the Ashkenazi community.[9,10] Thus, specific historical documents regarding the community here analysed were brought to shed light on this specific question providing new evidence to reconstruct a more detailed and accurate history of the Sephardic Jews.

Finally, Bedford and Yacobi raise a very interesting question based on the large excess of nonsynonymous over synonymous mutations. This has been detected previously in mtDNA phylogeny (younger mutations are enriched in the first type, so that they often define branch tips) and a lot of debate on the selective pressures acting upon is ongoing[11–16] (and was also detected in animal models as our group has shown in lab mouse[17]). The issue is, however, extraneous to the topic and goals of our paper; we can nevertheless clarify, as requested, that the finding does not result from a 'special clinical population of subjects'.

We certainly do agree that to know 'whether this is a coincidence, [or] a notable finding [...] would be helpful for evaluating the representativeness of [...] [our] results for Sephardim in Northeast Portugal and elsewhere'. Unfortunately, the answer to this alternative requires a much larger sampling effort and the cooperation of the scientific community devoted to demographic history of human maternal lineages as revealed by mtDNA.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Inês Nogueiro*[,1,2], João Teixeira[3], António Amorim[1,2], Leonor Gusmão[1,4] and Luis Alvarez[2]
[1]*Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP), Porto, Portugal;*
[2]*Faculty of Sciences, University of Porto, Porto, Portugal;*
[3]*Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany;*
[4]*DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Rio de Janeiro, Brazil*
E-mail: inogueiro@ipatimup.pt

1 Nogueiro I, Teixeira J, Amorim A, Gusmao L, Alvarez L: Echoes from Sepharad: signatures on the maternal gene pool of crypto-Jewish descendants. *Eur J Hum Genet* e-pub ahead of print 30 July 2014; doi: 10.1038/ejhg.2014.140.
2 Behar DM, Metspalu E, Kivisild T *et al*: Counting the founders: the matrilineal genetic ancestry of the Jewish Diaspora. *PLoS One* 2008; **3**: e2062.
3 Picornell A, Gomez-Barbeito L, Tomas C, Castro JA, Ramon MM: Mitochondrial DNA HVRI variation in Balearic populations. *Am J Phys Anthropol* 2005; **128**: 119–130.
4 Bedford FL. Sephardic signature in haplogroup T mitochondrial DNA. *Eur J Hum Genet* 2012; **20**: 441–448.
5 Bedford FL, Yacobi D, Felix G, Garza FM: Clarifying mitochondrial DNA subclades of T2e from mideast to Mexico. *J Phylogenet Evol Biol* 2013; **1**: 121.
6 van Oven M, Kayser M: Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* 2009; **30**: E386–E394.
7 Bedford FL, Yacobi D: On two Jewish clades in mitochondrial DNA. *Eur J Hum Genet* 2014; e-pub ahead of print 5 November 2014; doi:10.1038/ejhg.2014.231.
8 Torroni A, Huoponen K, Francalacci P *et al*: Classification of European mtDNAs from an analysis of three European populations. *Genetics* 1996; **144**: 1835–1850.
9 Roth C, Spivak A: *Historia de los marranos*. Buenos Aires: Editorial Israel, 1946.
10 Mea E: A Problemática do Judaísmo. Séculos XVI-XVII. *Congresso Internacional Inquisição Portuguesa. Tempo, Razão e Circunstância*. Lisboa-S: Paulo, 2007.
11 Mishmar D, Ruiz-Pesini E, Golik P *et al*: Natural selection shaped regional mtDNA variation in humans. *Proc Natl Acad Sci USA* 2003; **100**: 171–176.
12 Elson JL, Turnbull DM, Howell N: Comparative genomics and the evolution of human mitochondrial DNA: assessing the effects of selection. *Am J Hum Genet* 2004; **74**: 229–238.
13 Kivisild T, Shen P, Wall DP *et al*: The role of selection in the evolution of human mitochondrial genomes. *Genetics* 2006; **172**: 373–387.
14 Ruiz-Pesini E, Wallace DC: Evidence for adaptive selection acting on the tRNA and rRNA genes of human mitochondrial DNA. *Hum Mutat* 2006; **27**: 1072–1081.
15 Kimura M: The neutral theory of molecular evolution: a review of recent evidence. *Jpn J Genet* 1991; **66**: 367–386.
16 Soares P, Ermini L, Thomson N *et al*: Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am J Hum Genet* 2009; **84**: 740–759.
17 Goios A, Pereira L, Bogue M, Macaulay V, Amorim A: mtDNA phylogeny and evolution of laboratory mouse strains. *Genome Res* 2007; **17**: 293–298.

European Journal of Human Genetics (2015) 23, 993–995

www.nature.com/ejhg

LETTERS

On two Jewish clades in mitochondrial DNA

European Journal of Human Genetics (2015) **23,** 993–994; doi:10.1038/ejhg.2014.231; published online 5 November 2014

Recent research by Nogueiro *et al*[1] on mitochondrial DNA full genomic sequences from Northeast Portugal is a welcome addition to Sephardic genetics. We would like to clarify the published history of Sephardic T2e, a haplogroup found in three of their participants, in order to set the record straight. The researchers report on an exclusively Jewish signature within T2e that is characterized by a 9181 G mutation, apparently a new one based on their own observations of publically available sequences, their one new sequence, and PhyloTree. However, the data they draw from, as well as the conclusions they reach, were instead put forth by Bedford *et al*[2] in an article that is sited but for an unrelated reason.

Nogueiro *et al* state: 'Inside the T2e1 branch, a new sub-haplogroup T2e1b (Supplementary Figure 6), defined by the presence of variant m.9181 G, has been proposed in PhyloTree build 16.' The new branch and labels were not proposed by PhyloTree. They were proposed by Bedford *et al,* who report on 9181 G: 'The present 7 sequences along with the one on Genbank (Sephardic Bulgarian) and the pathological sequence from the medical literature were used to create a phylogenetic tree. The tree is presented in Figure 2. New labels for branches include T2e1b and T2e1b1 and are summarized for both this cluster and the previous 2308 G cluster in Table 2'. (As a historical note, Mannis van Oven, curator of PhyloTree, accepted our proposed branch assignments from Table 2 beginning with build 16; personal communication, email van Oven, 27 September 2013.)

Their article continues: 'T2e1b is supported by 11 complete mtDNA sequences, including the Bragança Jews. Except for two samples from mtDNA Community database, without information concerning their ethnicity, all the remaining nine individuals are Jews, Sephardim or Ashkenazim.' These ethnic observations, however, do not originate with Nogueiro *et al*, but instead from the work of Bedford *et al* who report in the Results section entitled 'The 9181 Cluster': 'For the total of 7 individuals with the defining mutation (kernel plus 6 matches), all 7 reported Jewish ancestry along the deep maternal line. This proved to be in sharp contrast to the 9 individuals without 9181 G, none of whom knew of any Jewish ancestry (Fisher exact test, $P<0.0001$). Details of the Jewish ancestry among the 9181 G group were 2 Sephardic (The Netherlands, Romania), 4 Ashkenazi (Czech Republic, Lithuania, Poland, unknown) and 1 unknown Jewish.' And again in the Discussion: 'Through the present full genomic sequencing of seven new samples, along with one on Genbank, and one in the medical literature, several striking aspects became apparent 1) All of those harboring 9181 G report Jewish maternal ancestry.'

The 11 complete mtDNA sequences referred to by Nogueiro *et al* include the single Genbank entry available before the study by Bedford *et al*. This was deposited by Behar *et al*[3] in connection with their finding that 9181 G within T2e (which they called 'T2f') was a founding lineage for Sephardic Bulgaria. It also includes the seven new mitochondrial full genomic sequences added to Genbank by Bedford *et al* in connection with their investigation of the mutation, the additional sequence summarized by Bedford *et al* from the medical literature, and the new one by Nogueiro *et al* from their current results, with identical motif to our samples from Sephardic Netherlands, Ashkenazi Poland, and Ashkenazi Czech Republic. (The 11th sequence from a publicly available database was one of our sequences that had not yet appeared on Genbank.)

The explicit *raison d'être* of Bedford *et al* was to investigate two T2e clusters with Sephardic Jewish affiliation, those with the 9181 G mutation and the signature we found earlier, 41C@-16114 T-16192 T-2308 G-15499 T. Thus, the above quoted passages from Bedford *et al* were not incidental or hidden within a different purpose. Nogueiro *et al* were also well acquainted with the details of the article: In an extended email enquiry, they asked about every sequence from our Figure 2 tree of 9181 G sequences, including the matching of submitted Genbank numbers to each node and the geographic locations and ethnicity we reported (8 email messages from Nogueiro to Bedford, 12 replies, 18–27 November 2013).

Near the conclusion of their article, Nogueiro *et al* restate that T2e is found in multiple Jewish groups and offer two alternative explanations: 'Remarkably, for two of the founder lineages (T2e1b and U2e1a), defined by the complete mitochondrial genome, the shared sequences belong to both Sephardic as well as Ashkenazi Jews.' We agree it is remarkable, hence our use of the term 'striking' (for T2e1b) as quoted earlier. They explain as follows: 'Two possible scenarios could accommodate this finding: either the defining variants for each branch could have arisen before the separation between the two Jewish groups; or there may have been recent introgression of Sephardic lineages into Ashkenazim communities in the north of Europe.' Likewise, they were our alternatives: 'The mutation is found in both those of Ashkenazi and of Sephardic maternal origin. Taken together, these findings suggest that this is a surprisingly old clade that may well predate the split between Jewish groups' and 'The present study also leaves open the possibility that the appearance of this mutation in both Sephardic and Ashkenazi Jewish groups is due to a more recent admixture between the Jewish groups rather than predating their split. For instance…'

Concerning our other Sephardic signature, which we labeled T2e1a1a, the work was acknowledged by Nogueiro *et al.* Although one of their three T2e results is a part of the T2e1b branch, their other two sequences fall into this T2e1a1a branch. To clarify potentially ambiguous phrasing about this signature in Nogueiro *et al* concerning the discovery of the back mutation at position 41, observations concerning 41 T/C instability, inferences of the ancestry of the Mexican samples as Sephardic, and establishment of higher branches from which the clade derives, these were all explicit in Bedford[4] and/or the Bedford *et al.* article. For instance, Nogueiro *et al* state, 'Considering the growing number of complete mtDNA sequences available at this time, it was possible to define a new sub-haplogroup T2e1a1, based on m.15499C>T variant.' In case it is unclear, we were the ones who contributed the growing number of sequences as well as explicitly defined T2e1a1 as a nesting structure headed by 15499 T and T2e1a by 2308 G.

The three full genomic sequence additions to T2e by Nogueiro *et al* reveal exciting new aspects of these two previously established Jewish T2e branches. For T2e1a1, they find both the first coding region private mutations and the first nested structure to be reported within this subclade. This suggests to us that estimates of age of this cluster[2,4]

be pushed back earlier. For T2e1b, they are the first to report its presence in Portugal. This suggests that, should the joint appearance in both Sephardim and Ashkenazim be due to a recent admixture, the direction of gene flow is more likely to have proceeded from Sepherad to Ashkenazi rather than the reverse. This was an issue concerning a recent admixture between the Jewish groups that was previously left unresolved.[2]

Finally and curiously, our perusal of the eight coding-region mutations described in the text of the article (Haplogroups HV0, N1, T2b, T2e, and U2) finds seven of eight of them to be nonsynonymous mutations, therefore altering the proteins manufactured from the DNA code and RNA translation. If the authors could clarify whether this is a coincidence, a notable finding, or reflects a special clinical population of subjects, it would be helpful for evaluating the representativeness of their results for Sephardim in Northeast Portugal and elsewhere.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Felice L Bedford* and Doron Yacobi
Department of Psychology, Cognitive Science, University of Arizona, Tucson, AZ, USA
E-mail: Bedford@u.arizona.edu

1. Nogueiro I, Teixeira J, Amorim A, Gusmão L, Alvarez L. Echoes from Sepharad: signatures on the maternal gene pool of crypto-Jewish descendants. *Eur J Hum Genet* 2015; **23**: 693–699.
2. Bedford FL, Yacobi, Felix G, Garza FM. Clarifying mitochondrial DNA subclades of T2e from Mideast to Mexico. *J Phylogenetics Evol Biol* 2013; **2**: 1–8.
3. Behar DM, Metspalu E, Kivisild T *et al*: Counting the founders: the matrilineal genetic ancestry of the Jewish Diaspora. *PLoS One* 2008; **3**: e2062.
4. Bedford FL. Sephardic signature in haplogroup T mitochondrial DNA. *Eur J Hum Genet* 2012; **20**: 441–448.

Research article:

Nogueiro, I., Teixeira, J. C., Amorim, A., Gusmão, L., & Alvarez, L. (2015). Portuguese crypto-Jews: the genetic heritage of a complex history. *Frontiers in genetics*, *6*.

frontiers in GENETICS

REVIEW ARTICLE
published: 02 February 2015
doi: 10.3389/fgene.2015.00012

Portuguese crypto-Jews: the genetic heritage of a complex history

*Inês Nogueiro[1,2,3] *, João C. Teixeira[4], António Amorim[1,2,3], Leonor Gusmão[1,3,5] and Luis Alvarez[1,2,3]*

[1] Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal
[2] Faculty of Sciences, University of Porto, Porto, Portugal
[3] Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal
[4] Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany
[5] DNA Diagnostic Laboratory, State University of Rio de Janeiro, Rio de Janeiro, Brazil

Edited by:
Eran Elhaik, University of Sheffield, UK

Reviewed by:
Eran Elhaik, University of Sheffield, UK
Luca Pagani, University of Cambridge, UK
Jill Bennett Gaieski, University of Pennsylvania, USA

***Correspondence:**
Inês Nogueiro, Institute of Molecular Pathology and Immunology of the University of Porto, Ipatimup, Rua Dr. Roberto Frias, s/n 4200-465 Porto, Portugal
e-mail: inogueiro@ipatimup.pt

The first documents mentioning Jewish people in Iberia are from the Visigothic period. It was also in this period that the first documented anti-Judaic persecution took place. Other episodes of persecution would happen again and again during the long troubled history of the Jewish people in Iberia and culminated with the Decrees of Expulsion and the establishment of the Inquisition: some Jews converted to Catholicism while others resisted and were forcedly baptized, becoming the first Iberian Crypto-Jews. In the 18th century the official discrimination and persecution carried out by the Inquisition ended and several Jewish communities emerged in Portugal. From a populational genetics point of view, the worldwide Diaspora of contemporary Jewish communities has been intensely studied. Nevertheless, very little information is available concerning Sephardic and Iberian Crypto-Jewish descendants. Data from the Iberian Peninsula, the original geographic source of Sephardic Jews, is limited to two populations in Portugal, Belmonte, and Bragança district, and the Chueta community from Mallorca. Belmonte was the first Jewish community studied for uniparental markers. The construction of a reference model for the history of the Portuguese Jewish communities, in which the genetic and classical historical data interplay dynamically, is still ongoing. Recently an enlarged sample covering a wide region in the Northeast Portugal was undertaken, allowing the genetic profiling of male and female lineages. A Jewish specific shared female lineage (HV0b) was detected between the community of Belmonte and Bragança. In contrast to what was previously described as a hallmark of the Portuguese Jews, an unexpectedly high polymorphism of lineages was found in Bragança, showing a surprising resistance to the erosion of genetic diversity typical of small-sized isolate populations, as well as signs of admixture with the Portuguese host population.

Keywords: crypto-Jews, Y chromosome, mtDNA, haplogroups, Portugal

INTRODUCTION

The purpose of this review is to summarize and critically revise the existing genetic data concerning the Portuguese Sephardic Jewish population. In this regard, other Sephardic population studies will be reviewed for contextualization. The historical background of the Sephardic Jews, with an emphasis on the Portuguese history, will be addressed. With the exception of the studies on the *Chuetas*, an isolated Mallorcan community from the Moslem period, 10–13th centuries (Santamaría Arández, 1997), currently available genetic studies on the original population from Iberia are restricted to Portugal, namely to Belmonte municipality and Bragança district (Northeast Portugal).

Until the decree of expulsion and the establishment of the Inquisition in Portugal during the 15th century, the Portuguese Jewish communities had quite a similar history. From the 15th century on, however, most of the Portuguese Jews were either exiled, or assimilated into the general population with the exception of a few Crypto-Jewish communities. The Crypto-Jewish phenomenon is defined as the secret adherence to Judaism while publicly professing another faith. These communities have kept, for more than 500 years, their hidden religious practices and their cultural identity using complex social strategies. Among these communities, special mention is due to Belmonte, a small town in the center of Portugal and also to several small villages and towns in the Bragança district. Uniparental genetic markers were typed in both communities (Adams et al., 2008; Nogueiro et al., 2010, 2014; Teixeira et al., 2011), showing genetic profiles and levels of genetic diversity in accordance with their dissimilar recent history compared to the Portuguese general population.

A TROUBLED HISTORY

The settlement of Jewish groups in Iberia undoubtedly occurred long time ago, although the exact date of their appearance is still today uncertain. The oldest archeological evidence of the Jewish presence in Iberia known so far, was recently found in the south of Portugal (Silves) with a chronology of 390 CE (http://www.uni-jena.de/en/News/PM120525_Schrifttafel.html).

Written documents mentioning the presence of Jewish communities in Iberia accumulate from the beginning of the Visigothic period onward: in the 4th century CE, the decisions of the Council of Elvira, particularly the interdiction of marriages between Jews and Christians, confirm their complete integration among Iberian communities (Martins, 2006; Wilke, 2009). The decree of expulsion of all Iberian Jews who would not embrace the Christian faith by the Visigothic King Sisebut in 613 CE marks the beginning of Crypto-Judaism and was a harbinger of their subsequent dramatic history in the Iberian Peninsula (Martins, 2006; Wilke, 2009).

A climate of tolerance between Jews, Muslims and Christians marked the Islamic period in Portugal (from 711 to 1249). Jews were at that time a demographically non-negligible minority with a very heterogeneous social status (Martins, 2006). Over the 11 and 12th centuries, a strong influence of Islamic culture was evident in the Iberian Peninsula, particularly in philosophy, geography, astronomy, mathematics and medicine, providing a cultural blossoming often qualified as the golden age of the Sephardic Jews (Maurício, 1971; Pignatelli, 2000; Tavares, 2004).

From the 12th to the end of the 13th century, the geography of Judaism changed significantly with the Christian settlement policies, to which the first Jewish medieval colonies owe their existence (Wilke, 2009). From the beginning of the Portuguese nation in 1143, until the Expulsion Edict in 1496, the successive Portuguese monarchs balanced the predominance of the Jewish social and economic life against the anti-Judaic clerical and popular pressures (Martins, 2006). Contrasting with the favorable measures toward the Jewish population, several opposite restrictive policies were also adopted by the successive Portuguese kings (Paulo, 1985; Azevedo, 1994; Pignatelli, 2000; Martins, 2006). The 14th century marks a dramatic change in this fragile equilibrium, which nevertheless allowed the emergence of Jewish communes in Portugal. The estimates of the size of the Portuguese Jewish population vary between 30 000 to more than 60 000. This number would increase substantially with the arrival of around 100 000 Spanish Jews fleeing from their country (Carvalho, 1999; Martins, 2006).

In 1478 the Inquisition was established in Spain, followed by the Edict of Expulsion by the Catholic Kings. These events would have a strong impact in the future of the Portuguese Jews (Canelo, 1987; Pignatelli, 2000; Martins, 2006). The initial tolerance toward the Spanish exiles and the Portuguese Jews was, however, doomed. The marriage of the Portuguese King Manuel I to the daughter of the Catholic Kings had severe political implications (Canelo, 1987; Pignatelli, 2000; Martins, 2006). Accordingly, in December of 1496 the King signed the Portuguese Edict of Expulsion, ordering the departure of Moors and Jews by October of the following year (Canelo, 1987; Pignatelli, 2000). However, in May of 1497, about 20 000 Jews from all over Portugal, who were preparing for exile, were forcibly baptized. As a result, there were officially no more Jews in Portugal, and instead, a new identity was created: the *new-Christians* or *Conversos* (Martins, 2006). The ambiguous policies adopted toward these communities, where some rulings favored the Jews and others worked against them, highlights their socioeconomic importance. In fact, King Manuel I of Portugal prohibited inquiries regarding the Jewish faith for 20 years, allowing for an accepted crypto-Judaism (Pignatelli, 2000; Martins, 2006).

The Papal Bull establishing the Inquisition in Portugal was issued in 1536 under the reign of King João III. In the 17 and 18th centuries, the inquisitorial processes intensified and as a result, there was a significant exodus of Jewish people to other countries, particularly of manufacturers and the merchant elite (Paulo, 1985).

By the end of the 15th century estimates of the size of the Jewish population were of about 100 000 people, which translates into 10% of all the Portuguese population at that time. The exact number of people who emigrated is not known, however, it is thought that in 1631 the Jewish population was reduced to 10 000 (Carvalho, 1999). Initially, the Portuguese Jews settled in Amsterdam, London, Hamburg, Turkey, some French and Italian cities, and North Africa; from the mid-16th century, they migrated to the Portuguese colonies in Africa, India and Brazil, and later from the several cities of northern Europe to the New World, (e.g., Curacao, Paramaribo, and the USA), where Jewish colonies founded synagogues with the Portuguese rite (Carvalho, 1999; Pignatelli, 2000; Martins, 2006).

The rebirth of the Jewish communities in Portugal took place in the early 19th century, when the Marquis of Pombal ended the official discrimination and persecution performed by the Inquisition (Carvalho, 1999; Martins, 2006). The Israeli Jewish community of Lisbon was founded by Sephardic Jews from North Africa (Pignatelli, 2000; Martins, 2006), and in the 20th century, the communities of Porto, Bragança, Belmonte, Faro and the Azores emerged (Paulo, 1985; Canelo, 1987; Pignatelli, 2000; Martins, 2006).

In particular, the communities of Bragança and Belmonte resurfaced thanks to the work of Samuel Schwartz and Captain Barros Basto, who started a movement in the early 20th century which aimed to bring back the Crypto-Jews to normative Judaism. While in Belmonte the community is still dynamic today, the Bragança community was dissolved in 1934, shortly after its appearance and its population dispersed in the region, though a strong sense of identity among their Jewish descendants is still well alive today.

THE CRYPTO-JEWS' GENETIC HERITAGE

The genetic heritage of Jewish populations has been deeply scrutinized at the population level as well as for the medical implications, using uniparental and autosomal markers (Hammer et al., 2000; Ostrer, 2001; Bauchet et al., 2007; Adams et al., 2008; Behar et al., 2008; Olshen et al., 2008; Kopelman et al., 2009; Elhaik, 2013) and more recently through genome-wide approaches (Seldin et al., 2006; Atzmon et al., 2010; Behar et al., 2010; Campbell et al., 2012; Velez et al., 2012; Ostrer and Skorecki, 2013).

In the medical field, the Ashkenazi community has been by far the most investigated Jewish group. There is a vast list of published studies focusing on genetic diseases in the Ashkenazi population (Ostrer, 2001; Alcalay et al., 2014; Feldman et al., 2014; Tafe et al., 2015) contrasting with the low number of such studies in other Jewish groups, particularly in the Sephardic group. The term "Sephardic" frequently includes not just the original Portuguese/Spanish Jewish populations but also all other Iberian exiled communities that follow the Sephardic rite and it can as well be used as a synonym for non-Ashkenazi groups, oftentimes

engulfing the Mizrahi group. Thus, genetic diseases commonly found in Sephardic Jews can comprise particular disorders that are exclusive to a specific sub-populations of this heterogeneous group (Rosner et al., 2009).

With regard to the genetic disorders of the Sephardic Jews who stayed in Iberia after the decrees of expulsion and the establishment of the Inquisition, very little is known. The only recognized Jewish population in the Spanish territory that follows this criterion are the *Chuetas*. Several reports are available with clinical relevance (Buades et al., 1995; Domingo et al., 2000; Guix et al., 2002; Cambra et al., 2009). As to Portugal, there is only one report on a Jewish genetic condition, an autosomal recessive form of retinitis pigmentosa studied in the Crypto-Jews of Belmonte (Gerber et al., 2000). Similar to other Sephardic specific variants, as the consequence of sustained inbreeding practices, this seems to have arisen 200–500 years ago, after the establishment of this isolated population in Belmonte (Gerber et al., 2000). Recent studies also reported a high prevalence of this disorder in Ashkenazi Jews (Zelinger et al., 2011; Zuchner et al., 2011; Venturini et al., 2014) but caused by a mutation in a different gene.

MONOPARENTAL GENETIC MARKERS

In recent years, the analysis of uniparentally inherited genetic systems, the non-recombining region of Y chromosome (NRY) and mitochondrial DNA (mtDNA), has played a central role in disclosing the demographic events that have shaped modern human population structure (Brown, 1980; Cann et al., 1987; Hammer, 1994; Jobling and Tyler-Smith, 2003). Both genetic systems have been used since the 90 s in the analysis of Iberian Peninsula populations, producing a detailed genetic landscape (Arroyo-Pardo et al., 2007; Adams et al., 2008; Capelli et al., 2009; Santos et al., 2014) and the assessment of the contributions from the various parental populations to the Iberian genetic composition is currently a very active research topic.

Y CHROMOSOME STUDIES IN SEPHARDIC POPULATIONS

The search for a common Middle Eastern origin of contemporary Jewish male lineages started with Hammer and Skorecki's landmark genetic study of the Cohanim (Hammer et al., 1997), a priestly lineage of the Jewish religion. In this study they defined the "Cohen Modal Haplotype" or "CMH" and showed a common origin for this lineage among both Sephardim as well as Ashkenazim Jews. Thomas et al. (1998) found that Y chromosomes of present-day Cohanim and Levites (also a priestly lineage) shared a common origin estimated to date about 3000 years before present.

Subsequent studies (Hammer et al., 2000; Thomas et al., 2000; Nebel et al., 2001) have suggested that most Jewish communities have remained quite isolated from neighboring non-Jewish communities during and after the Diaspora and that the communities from Europe, North Africa, and the Middle East descended from a common Middle Eastern ancestral population.

A high-resolution Y chromosome haplotype analysis on unrelated Israeli and Palestinian Moslem Arabs showed a common pool for the male lineages. However, some significant differences were also detected between Jews and Arabs, suggesting a recent divergence of the Arab clade from the common ancestral population (Nebel et al., 2000).

The research on the Jewish Priestly lineages, Levites and Cohanim was again addressed by Behar et al. (2003) showing that paternal ancestries of Ashkenazi and Sephardi Levites are genetically dissimilar, in contrast to what was found for Ashkenazi and Sephardic Cohanim (Hammer et al., 2009).

A very recent study (Tofanelli et al., 2014) of the haplotype motifs of Levites and Cohanim Jewish Priestly lineages has, however, found that these supposed markers of Jewish ancestry can lead to ambiguous results since they are not identical by descent. These motifs were observed in independent lineages from different ethnic, cultural and geographic groups, probably due to multiple founder events, recombination and admixture of the Jewish genetic pool in the course of their history (Tofanelli et al., 2014).

Sephardic Jews were also included in studies of the Y chromosome phylogeography (Semino et al., 2004), in the genetic affinities between Jews and other populations from the Middle East (Shen et al., 2004; Oefner et al., 2013) and in the construction of the genetic landscape of the Iberian Peninsula (Adams et al., 2008). The genetic profile of the Sephardic Jews can be seen in **Figure 1**.

Several studies refer to the putative contribution of Sephardic Jews to the Y chromosome genetic pool of particular geographic regions, like the Azores and Madeira islands (Goncalves et al., 2005; Pacheco et al., 2005) and also to the New-World, such as Brazil (Carvalho-Silva et al., 2001) or New Mexico and southern Colorado (Sutton et al., 2006), where many populations settled, including migrants from Portugal/Iberia.

Y CHROMOSOME IN SEPHARDIC PORTUGUESE JEWS

The profile of male lineages in Portugal was drafted in a study comprising 663 male samples from the 18 administrative districts of Portugal and a typical western European composition was demonstrated by the high frequencies of haplogroups R1b1a-M269 (57.7%), I-M170 (6.1%), G-M201 (5.5%), and E1b1b-M81 (5.6%), as well as a Middle Eastern influence, denoted by the presence of J-12f2.1 lineage (10.4%; Beleza et al., 2006). Possible Sephardic contributions to this genetic pool were also addressed in some reports (Goncalves et al., 2005; Pacheco et al., 2005) but very little was known about the Portuguese Jews, even though, in a large scale study of the Iberian genetic diversity, very few Jewish male samples from Belmonte were analyzed (Adams et al., 2008).

These samples were included in a larger group of self-defined Sephardic Jewish males not only from the Iberia Peninsula but also from other countries that received Jewish exiles after the decrees of expulsion in the 15th century. This group of self-defined Sephardic Jews was treated as a single group, therefore the inference of a genetic profile for the Portuguese Jews was not possible from the published data. In the following, we will focus on the results obtained exclusively for the 16 Jewish samples of Belmonte, obtained upon request to the authors (Adams et al., 2008).

The genetic profile of the Portuguese Jewish and non-Jewish male lineages can be seen in **Figure 2**. The Y chromosome SNPs analyzed allowed the definition of just three different lineages in Belmonte Jews: eleven individuals were classified as J-12f2.1, four as R1b1a-M269 and one as G-M201, with a frequency of 68.8, 25 and 6.2% respectively.

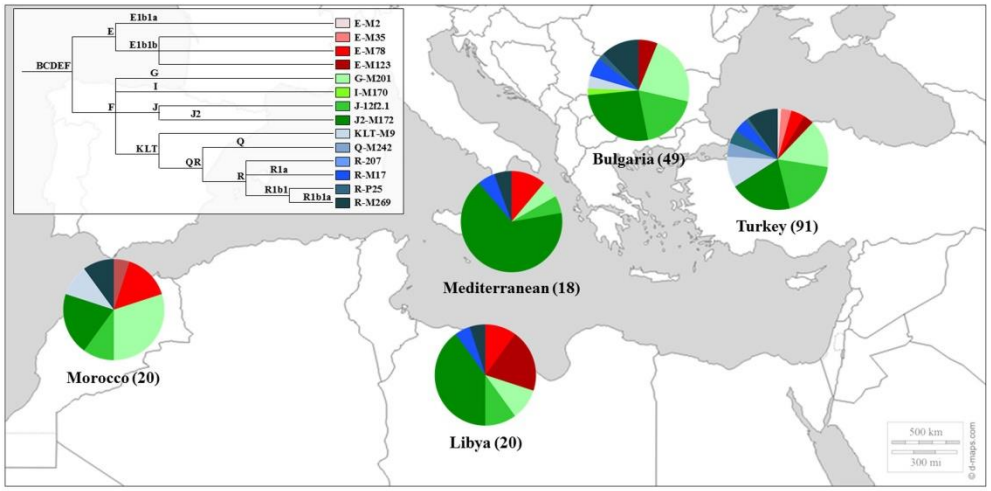


FIGURE 1 | Y Chromosome haplogroup distributions of Sephardic Jewish populations. Sectors in pie charts are proportional to haplogroup frequency. Number of total individuals (n) are in brackets for each population. Jewish samples are from the works of Shen et al. (2004) and Adams et al. (2008).

The analyses of the STRs DYS19, DYS388, DYS389I-II, DYS390, DYS391, DYS392, DYS393, DYS438, DYS439, DYS385a, and DYS385b revealed a total of only four distinct haplotypes. In the R1b1a-M269 haplogroup two different haplotypes were detected, diverging one from the other by one mutation step (DYS389II), inside the J-12f2.1 haplogroup all the eleven individuals presented exactly the same haplotype, reflecting very low levels of genetic diversity among this Jewish community.

A completely different picture of the Portuguese male Jewish lineages was, however, brought to light when the descendants of the crypto-Jews from Bragança district were analyzed (Nogueiro et al., 2010). In this study, 57 unrelated self-designated Jewish males from the Northeast Portugal (Bragança, Argozelo, Carção, Mogadouro, and Vilarinho dos Galegos) were selected, using a combination of geographic, religious ethno-historical and affiliation criteria.

The SNPs typed allowed the discrimination of 10 different haplogroups and the analysis of the Y-STR loci revealed 41 different haplotypes. The most frequent haplogroups found were R1b1a-M269, J-12f2.1, and T-M70, adding up to 80.7% of the total sample (**Figure 2**).

The effect of genetic drift in an isolated, small sized population could explain the high frequency found in Bragança for lineages typically predominant in other Jewish populations, such as J-12f2.1 (36.8%) and T-M70 (15.8%). However, the high haplogroup diversity combined with the high (intra-haplogroup) haplotypic diversity are extremely surprising, as they show exactly the opposite of what is expected, namely a deep genetic diversity loss. Although inbreeding practices were sustained by the Portuguese crypto-Jewish communities, in the light of the obtained results it seems that its effects were less pronounced in the Bragança district compared to Belmonte, due probably to complex mating strategies and/or a very heterogeneous genetic pool in their origin.

Haplogroup R1b1a-M269, representative of Western Europe, is the most common lineage found in the Portuguese general population (57.7%; Beleza et al., 2006) while in the Jewish population it does not go beyond 28.1%. In contrast, haplogroup J-12f2.1, typical in other Jewish populations, appeared with a frequency of 36.8%, contrasting with the low frequency found in the Portuguese population (10.4%; Beleza et al., 2006). The same happened with haplogroup T-M70 (15.8%), which is quite rare among the non-Jewish Portuguese lineages (1.6%; Beleza et al., 2006).

Haplogroup R1b1a-M269, emerges as the most frequent lineage in European individuals. Its distribution displays an increasing gradient moving from east to west (Semino et al., 2000; Bosch et al., 2001; Cruciani et al., 2002; Flores et al., 2004; Brion et al., 2005; Moore et al., 2006), being the most frequent haplogroup in the Iberian Peninsula. Lineage R1b1a-M269 was associated with the expanding Neolithic movements from the Near East to the western fringe of Europe, although this is still a matter of debate (Balaresque et al., 2010; Busby et al., 2011).This haplogroup was absent in Jewish population studies until the report of Adams et al. (2008).

The high frequency of haplogroup R1b1a-M269 found in both groups of Portuguese Jews could result from admixture with the Portuguese/Iberian population, and/or from introgression before their entry into the Iberian Peninsula. Pairwise R_{ST} genetic distances between R1b1a-M269 Jewish haplotypes from Bragança district and those from Portugal (Beleza et al., 2006) and Turkey (Cinnioğlu et al., 2004), were analyzed by Nogueiro et al. (2010) to

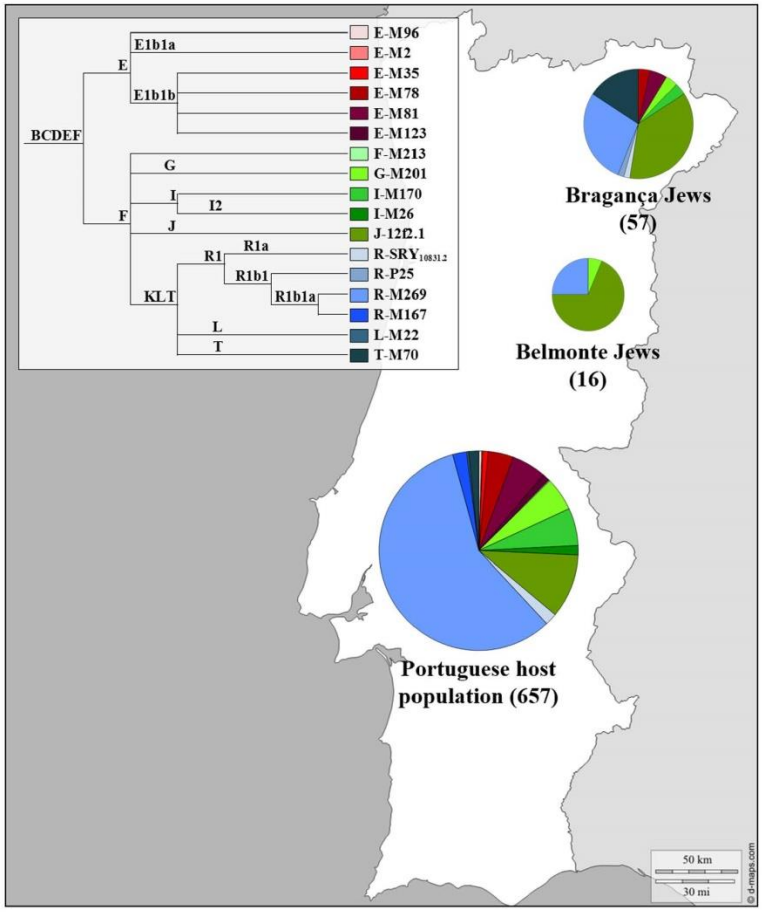


FIGURE 2 | Y Chromosome haplogroup distributions of the Portuguese Sephardic Jews and non-Jewish population. Sectors in pie charts are proportional to haplogroup frequency. Number of total individuals (n) are in brackets for each population. Jewish samples (Belmonte and Bragança) are from the works of Adams et al. (2008) and Nogueiro et al. (2010), respectively, and the Portuguese host population from Beleza et al. (2006).

verify the contributions of Western Europe versus Near East to the frequency of the Portuguese Jewish R1b1a-M269. A lower genetic distance between the Portuguese Jewish and non-Jewish R1b1a-M269 haplotypes than between these two samples and the one from Turkey was detected. Thus, an important Western European R1b1a-M269 introgression into the Portuguese Sephardic Jews, most probably after their arrival in Iberia, is the most plausible scenario (Nogueiro et al., 2010).

No exact matches, or total identity in all markers, were found between R1b1a-M269 haplotypes from Bragança district and those from Belmonte Jews. However, only two mutation steps distinguish each of the two Belmonte haplotypes (which differ by just one mutation step) from the one found in Bragança.

Haplogroup J-12f2.1 has a Middle Eastern origin and includes two groups, the J1-M267 and the J2-M172. Lineage J2-M172 is more common and is widely spread over Europe, particularly in the Mediterranean basin (Semino et al., 2004). Haplogroup J-12f2.1 presents a decreasing gradient from its origin toward Europe and is associated with the demic diffusion of the Neolithic farmers (Underhill et al., 2001; Semino et al., 2004) and also to more recent events, such as the Phoenician maritime migrations along the Mediterranean (Hammer et al., 2000; Di Giacomo et al., 2004; Zalloua et al., 2008).

This haplogroup is referred to as being predominant in diverse referenced Jewish populations (Hammer et al., 1997, 2000; Nebel et al., 2001; Adams et al., 2008), reaching in Sephardic Jews,

frequencies above 40% (Semino et al., 2004; Adams et al., 2008). While in Portugal it accounts for 3.4% of J-12f2.1 and 7% of J2-M172 lineages (Beleza et al., 2005) in the Portuguese Jews, it reached values of 68.2% for J-12f2.1 in Belmonte, 12.3% and 24.5% for J1-M267 and J2-M172 respectively, in the Bragança district.

The high frequencies of J-12f2.1 haplogroup found in both groups of Portuguese Jews, compared to the non-Jewish Portuguese host population, could therefore represent part of the genetic pool of the ancestral Sephardic population that established the first Jewish settlements in Portugal. For this lineage, again, no exact matches were found between the haplotypes of Belmonte and Bragança district. Three mutational steps apart, one haplotype of Belmonte matches five individuals from Bragança, Carção, Argozelo, and Vilarinho dos Galegos.

The presence of the mutation M70 defines haplogroup T. Its origin is attributed to the Middle East (Underhill et al., 2001) and from there it spread along the Mediterranean and East Africa. It is a rather rare haplogroup, displaying a global frequency of around 1% (King et al., 2007), but nonetheless it is found at quite high frequencies in Sephardic Levites (23%) and Sephardic Israelis (13%; Behar et al., 2004).

In Portugal it accounts for just 1.6% (Beleza et al., 2006) but reaches 15.8% in Bragança district Jews, being absent in the Belmonte samples. This lineage probably represents a relic of the original Sephardic male genetic pool, since it appears with similar frequencies in Israeli Sephardic Jews, but is quite rare in the Mediterranean coast and in Iberia.

Several other haplogroups were detected in the NE Portuguese Jews with residual frequencies, namely E1b1b-M78 with 3.5%, E1b1b-M81 with 5.2%, I-M170with 3.5% R1b1-P25 with 1.8%, R1a-SRY$_{10831.2}$ with 1.8%, and G-M201 with 3.5%.

G-M201 was also detected in Belmonte (6.2%) at about the same frequency as in the non-Jewish Portuguese population. Adams et al. (2008) suggested that this haplogroup could reveal an introgression of Sephardic Jews into the Iberian population. However, the estimated age for this lineage in Portugal (Beleza et al., 2005) is consistent with its introduction during the Neolithic and the results of relative frequencies and STR variance inside this lineage from Adams et al. (2008) does not allow the definition of the gene flow direction.

mtDNA STUDIES IN SEPHARDIC POPULATIONS

The first work dealing with the maternal (mtDNA) lineages in Sephardic populations dates from 1986 (Bonne-Tamir et al., 1986). The authors studied mtDNA variation patterns in a sample of 81 Arab and Jewish Israelis, including three individuals of Sephardic origin and a possible existence of group-specific mtDNA fragment patterns was speculated. Shortly after, in 1991, a complementary work (Tikochinski et al., 1991) increased the available data, with 39 Jewish individuals including 18 Sephardic samples, mainly of Moroccan origin. Twenty-one distinct maternal linages were identified but no estimation of the introgression degree from the host population was performed.

Later (Thomas et al., 2002), HVRI segments of the mtDNA Control Region (CR) were sequenced in 615 Jewish individuals belonging to nine geographically separated groups. The work included a large sample of 115 Moroccan Jews, a community that, as previously stated, received Iberian Jews after their expulsion. The HVRI frequencies for the Moroccan Jews showed a high prevalence (27%) of sequences presenting no differences to the revised Cambridge Reference Sequence (rCRS), and so, included in the H haplogroup, which although ubiquitous in Europe has a significantly higher prevalence in the Iberian Peninsula.

Similar results were found in later works (Picornell et al., 2006; Behar et al., 2008), including samples of Sephardic Jewish communities. In the work of Picornell et al. (2006) the HVRI and HVRII were analyzed in 31 individuals from Turkey and 12 from Morocco while Behar et al. (2008) studied a larger fragment of the mtDNA control region (16,024-300) together with diagnostic positions of the coding region in 149 Moroccans and 213 individuals from different Sephardic communities (e.g., Bulgarian and Turks). These communities also presented a high proportions of mtDNA haplogroup H. Moreover, a remarkable west Mediterranean imprint among the Turkic Jewish sample was observed (Behar et al., 2008). The genetic profile of the referred Jewish's populations is shown in **Figure 3**.

The detection of traces of Sephardic Jewish presence was addressed in several populations through the analysis of their possible contribution to the genetic background of the host populations. Such contribution was investigated in regions like the Portuguese north Atlantic archipelagos of Madeira and Azores (Brehm et al., 2003; Santos et al., 2003, 2010) or the Spanish population of Pasiegos, an isolate from the northern region of Cantabria (Maca-Meyer et al., 2003). HVRI was also analyzed in a sample of 45 *Chuetas* (Picornell et al., 2005) showing a high frequency (23% the most prevalent) of Middle Eastern haplogroup R0a. This pattern of mtDNA diversity, showing haplogroups from the Middle East, was associated with female-specific founder events and has been described in various Jewish communities (Thomas et al., 2002; Behar et al., 2008).

Such contribution was also analyzed in the New World populations, since Jewish migrations after the expulsion are well documented (Barnavi et al., 1992; Bacon, 2011). One of the first works was produced by Carvajal-Carmona et al. (2000), comprising 80 Colombian individuals from the Antioquia province. Only restriction diagnostic sites of the four major founder Native American mitochondrial haplogroups (A–D) were examined, and therefore, due to the low resolution power, it was not possible to carry out further insights on the ancestry of the non-Amerindian lineages detected (~10%).

Two Latin American populations, (inhabitants of Loja province in Sothern Ecuador and the Hispanos community from San Luis Valley of Colorado in Southwest US), with a presumable Sephardic Jewish ancestry were analyzed by Velez et al. (2012). The selection of such communities was done based on the occurrence of two mutations previously associated to genetic conditions described in different Jewish communities (Ostrer, 2001). HVRI mtDNA sequence variation and RLFP analysis of diagnostic haplogroup positions of the coding region from 53 individuals revealed, as in the previous work, a high prevalence of the founder Native American lineages (~93%).

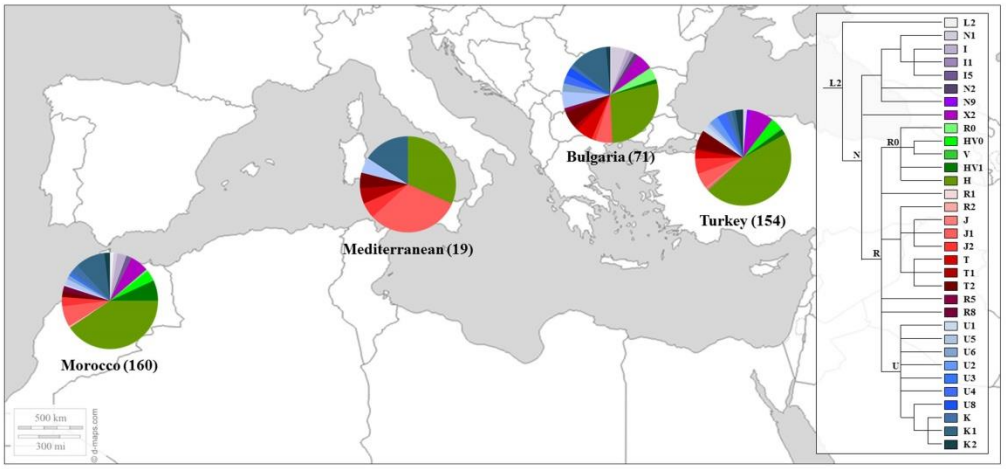


FIGURE 3 | Mitochondrial haplogroup distributions of the Sephardic Jewish populations. Sectors in pie charts are proportional to haplogroup frequency. Number of total individuals (n) are in brackets for each population. Jewish samples are from the works of Picornell et al. (2006) and Behar et al. (2008).

Complementary strategies have also been used to study this possible Sephardic contribution in a more recent work (Bedford, 2011). The analysis of CR shared haplotypes in T2e haplogroup, using different public mtDNA databases, was reported. The authors focused the analysis in a specific CR motif (combination of certain variations) initially named as T2e5, which can be currently located into the T2e1a1a1 clade, according to the updated mtDNA phylogeny in PhyloTree built 16 (van Oven and Kayser, 2009). The complete mtDNA genome from a Turkish Sephardic individual belonging to this rare clade was also sequenced. The authors observed the motif in twelve samples including Sephardic descendants from Turkey and Bulgaria, individuals from North American regions (Northern Mexico and South USA, places known for receiving Spanish *Conversos*), and samples from Portugal and Brazil, also consistent with a speculated Sephardic ancestry.

Further insights regarding the Sephardic signature inside T2e haplogroup and the genetic affinities of the T2e Northern Mexican samples were presented later (Bedford et al., 2013). The mitogenomes analyzed allowed to clarify the phylogeny of the Sephardic branches T2e1a and T2e1b. Indeed, the perfect match between complete sequences of Mexican individuals belonging to T2e1a clade and those from Turkish/Bulgarian Sephardic individuals, provided genetic evidence for a Sephardic origin of this lineage.

mtDNA IN SEPHARDIC PORTUGUESE JEWS

In parallel with the analysis of the male counterpart (Y chromosome), the mtDNA variation in Portugal was used to investigate the maternal heritage in the current Portuguese genetic landscape. The first detailed report on the Portuguese mtDNA, was done by Pereira et al. (2000), where HVRI and HVRII of 549 samples from North, Central and South Portugal were typed. The Portuguese population presented a typically Western European mtDNA composition with the distinction of harboring higher frequencies of North and Sub-Saharan African specific lineages (haplogroups L1-3 and U6). Similar results were found in a later work (Gonzalez et al., 2003), with the sequencing of the control region HVRI in 299 Portuguese samples. Specific areas of Portugal were analyzed in detail, due to their distinctive geographic and demographic characteristics: Azores and Madeira islands have a recent settlement history and played an important role in the modern slave trade from Africa to the New World, which is reflected in the significant presence of sub Saharan lineages (Brehm et al., 2003; Santos et al., 2003, 2010); The populations of Coruche, Pias, and Alcacer do Sal, were analyzed due to a recent malaria endemicity with different mtDNA compositions (Pereira et al., 2010); The Northeast population of Miranda do Douro near the Portuguese-Spanish border (Mairal et al., 2013) was also recently studied for the singularity of their Mirandese language and a diversity decrease of mtDNA lineages was found for this small-sized and isolated population.

The possible contribution of Sephardic lineages to the female Portuguese genetic pool was investigated as previously stated (Brehm et al., 2003; Santos et al., 2003, 2010) nevertheless only three published reports on mtDNA variation actually included Portuguese Sephardic Jewish descendants (Behar et al., 2008; Teixeira et al., 2011; Nogueiro et al., 2014). Behar et al. (2008) analyzed mtDNA CR along with diagnostic positions of the coding region in 30 Jewish individuals from the community of Belmonte municipality in Central Portugal, while the most recent one, studied the complete mtDNA molecule in 57 self-designated Jewish descendants, sampled in several locations in Bragança district (NE Portugal): Bragança, Argozelo, Carção, Mogadouro and Vilarinho

dos Galegos, (Nogueiro et al., 2014). The genetic profile of the Portuguese Jews and the Portuguese non-Jewish female lineages can be seen in **Figure 4**.

The Belmonte community presented a very low diversity, with only two lineages detected, and all samples inside each haplogroup presented the same haplotype. The distribution of the haplogroups is uneven, since one of them, HV0b haplogroup, stands out with a frequency of 93.3%. On the other hand, the Bragança Jewish sample presented a much higher haplogroup diversity (32 haplogroups) as well as intra-haplogroup diversity (45 haplotypes). The haplogroups that stand out in the Bragança Jewish sample, when compared with the Portuguese host population, were: HV0b, N1, T2 and U2, adding up to 38.6%.

The single mtDNA lineage shared between the two analyzed Portuguese Jewish series, Belmonte and Bragança, is HV0b. The expansion of the haplogroup HV0, including the haplogroup V, had been proposed to have an Iberian origin, after the last glacial maximum, and nowadays can be also found in North African populations. However, little is known about the distribution of the HV0b clade. According to the CR mutations, the specific defining combination variants (16298C, 72C, 195C, 198T and 263G) can be found at EMPOP forensic database in nine individuals, including one from Northern Africa; five from Central and North America; two from Central Europe; and one from Southern Portugal. Moreover, a recent paper focused on mtDNA variation in Andalusia (Southern Spain), (Hernández et al., 2014) reported samples that can be included inside this haplogroup but with additional variations: one sample from Granada (16093C, 16293G, 16298C, 16519C, 72C, 195C, 198T, and 263G) and two from Huelva (16153A, 16298C, 72C, 195C, 198T, and 263G), the latter being also found in ethnic Portuguese Roma (Mendizabal et al., 2011). This scenario is compatible with its origin in the Iberian Peninsula and with an early introgression into the Iberian Jewry gene pool from its host population, as suggested by Behar et al. (2008).

Looking deeper into the variability of the HV0b clade through the comparison of complete mtDNA sequences from Bragança Jews (seven), Belmonte Jews (one) and GeneBank (two), reveals a shared common private variant in the coding region (8520G) of the Portuguese Jewish samples. Moreover, the Bragança samples cluster together, sharing a variant (10644A), which seems to have arisen locally. Although new complete mtDNA sequences are needed for a better understanding of this particular clade, the results available so far support the hypothesis that at least the HV0b-8520G haplotype is a Sephardic Jewish founding lineage.

Another interesting clade that deserves further attention is T2e1. This clade was described as one of the founder lineages of the Bulgarian Sephardic community (Behar et al., 2008), and as referred to previously, is found among populations from Northern Mexico and Southern USA, being interpreted as a Sephardic signature (Bedford, 2011; Bedford et al., 2013). A Near East origin for haplogroup T was proposed with a posterior expansion into Europe before the Neolithic. The distribution of some sub-clades, including the T2e lineage, was associated with posterior European indigenous dispersion events (Pala et al., 2012). Three Bragança Jewish samples belong to T2e1 clade, two of which can be further classified as T2e1a and one as T2e1b, according to coding region variations (2308G and 9181G, respectively; van Oven and Kayser, 2009). Inside T2e1a branch the Bragança Jewish samples can be further classified into T2e1a1 sub-branch, based on two CR variations (back-mutation at position 41 and 16192T) and as T2e1a1a based on the presence of the CR variation (16114T). Inside T2e1a1a, Bragança Jewish sequences cluster together with three samples, one Sephardic sample from Turkey, and two other from Northern Mexico and Southern USA (Bedford et al., 2013). Moreover, the Braganca Jews present two further distinct coding region variants (13135A and 7133T). As stated by Bedford et al. (2013), the ancestry of the analyzed Northern American samples are consistent with an Iberian Sephardic origin. Apart from the Bragança T2e1b sample, this branch contains 10 other complete mtDNA sequences. Eight of them, with widespread geographic European origin, are of Jewish ethnicity, either Sephardic (Netherlands, Romania and Bulgaria) or Ashkenazi (Poland, Czech Republic, and Lithuania; Bedford et al., 2013).

Similar results regarding shared haplotypes between samples from Sephardic and Ashkenazi origin were found inside U2e1a1 sub-clade. All four U2e1a1 samples from Bragança Jews share the same haplotype with a Jewish Ashkenazi sample from Moldova (Behar et al., 2012), presenting two private CR variations (8014G and 13708A).

This pattern of shared haplotypes between Sephardim and Ashkenazim samples, firstly described by Bedford et al. (2013), could represent two possible scenarios (Nogueiro et al., 2014): the defining variants could have arisen before the separation between the two Jewish communities; or it may have resulted from a recent introgression of Sephardic lineages into the Ashkenazi gene pool. Further genetic data will help to clarify this issue, but it is possible to add non-genetic evidence for the second hypothesis, since marriages between members of the two communities have been recorded (Roth, 1979; Elvira, 2007), the descendants having been assimilated into the Ashkenazi community.

FINAL REMARKS

In conclusion, the demographic processes underlying the genetic pool of the Portuguese Crypto-Jews descendants studied so far, are much more complex than would be expected under the classical model of extreme inbreeding and drift, with consequent loss of genetic diversity. The contrasting patterns observed in Bragança community and Belmonte are enough to sustain that whatever the results of future studies, no simple and uniform evolutionary model will accommodate the sharp heterogeneity already observed.

Notwithstanding this difference, both groups display a genetic pool clearly showing contributions of European and Near Eastern lineages, in accordance with a significant persistence of a Jewish heritage, translated in a conscience of belonging to a distinctive community. This ancestry was detected within both male and female lineages, indicating that introgression from and admixture with the host population does seem to have been significantly gender biased.

Moreover, the high genetic diversity found in Bragança demonstrates that there was neither a low number of founder lineages, nor a significant reduction of effective population size as indeed occurred in Belmonte. It remains to be explained how this

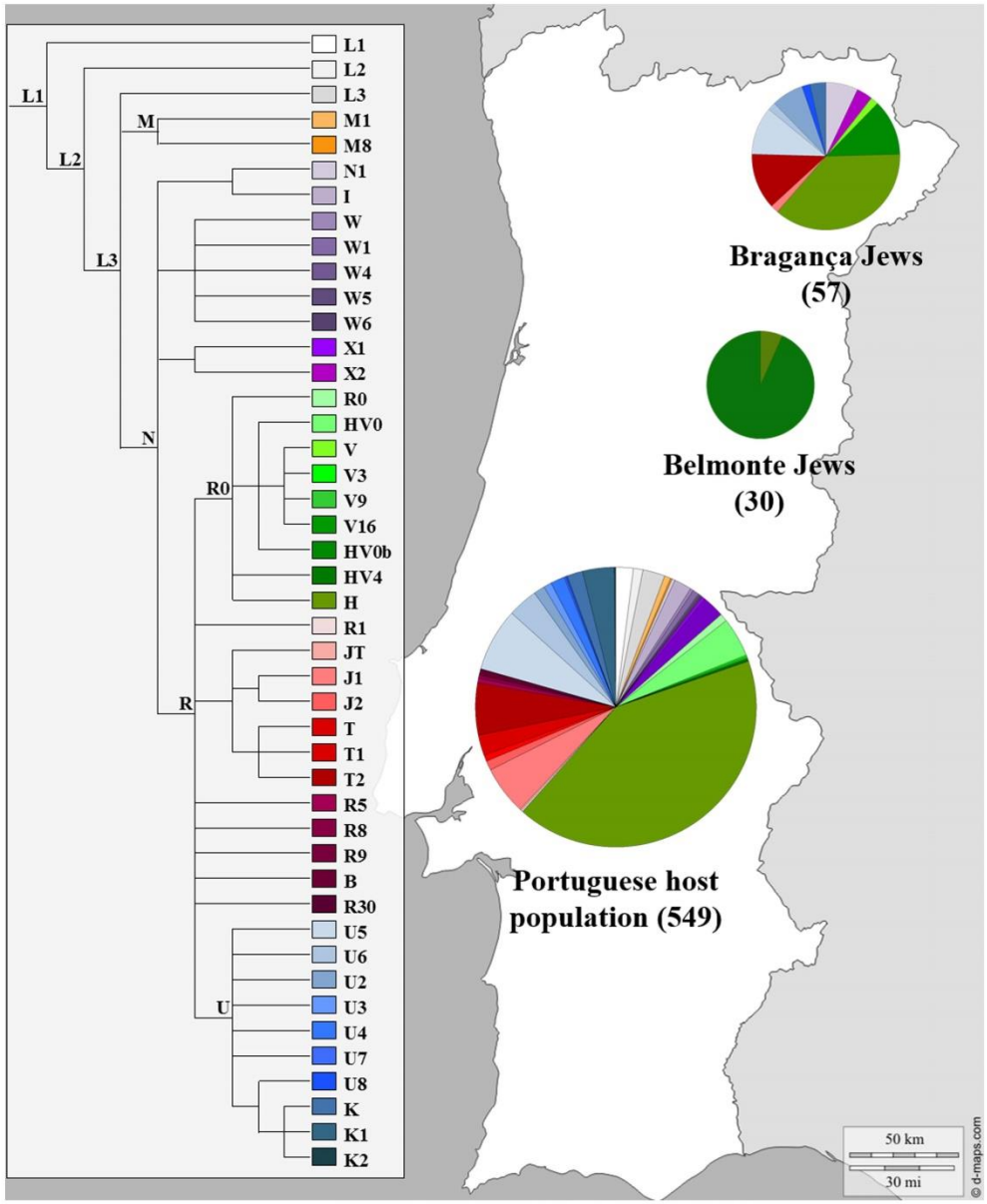


FIGURE 4 | Mitochondrial haplogroup distributions of the Portuguese Sephardic Jews and non-Jewish population. Sectors in pie charts are proportional to haplogroup frequency. Number of total individuals (n) are in brackets for each population. Jewish samples (Belmonte and Bragança) are from the works of Behar et al. (2008) and Nogueiro et al. (2014), respectively, and the Portuguese host population from Pereira et al. (2000).

resistance to genetic erosion, as expected in endogamous, small sized populations, was achieved, that is to say, what mating strategies were undertaken by these communities which ensured a steady gene flow between them, counteracting the expected inbreeding.

New data from recombining genetic markers in the line of Behar et al. (2010), as well as from classical genealogical studies, will surely contribute decisively to explain how this was achieved. At any rate, the DNA evidence gathered so far adds a new facet to the already recognized astonishing cultural resistance of these communities: not only they have kept a sense of belonging throughout centuries of persecution but they also succeeded in maintaining a genetic heritage of their own.

ACKNOWLEDGMENTS

We thank Doron Behar and Mark Jobling for the additional information concerning the Portuguese Jewish samples from Belmonte, as well as all DNA donors for their collaboration and generosity and Kevin Charles Rowe for the English language revision. This work was supported by FCT, Portuguese Foundation for Science and Technology, financed by the European Social Funds (COMPETE-FEDER), and project PTDC/ATP-DEM/4545/2012. Luis Alvarez (SFRH/BPD/65000/2009), Inês Nogueiro (SFRH/BD/73336/2010) and João C. Teixeira (SFRH/BD/77043/2011) are supported by FCT fellowships and funded by POPH-QREN – Promotion of scientific employment, supported by the European Social Fund and national funds of the Ministry of Education and Science. IPATIMUP is an Associate Laboratory of the Portuguese Ministry of Science, Technology and Higher Education and is partially supported by FCT.

REFERENCES

Adams, S. M., Bosch, E., Balaresque, P. L., Ballereau, S. J., Lee, A. C., Arroyo, E., et al. (2008). The genetic legacy of religious diversity and intolerance: paternal lineages of Christians, Jews, and Muslims in the Iberian Peninsula. *Am. J. Hum. Genet.* 83, 725–736. doi: 10.1016/j.ajhg.2008.11.007

Alcalay, R. N., Dinur, T., Quinn, T., Sakanaka, K., Levy, O., Waters, C., et al. (2014). Comparison of Parkinson risk in Ashkenazi Jewish patients with Gaucher disease and GBA heterozygotes. *JAMA Neurol.* 71, 752–757. doi: 10.1001/jamaneurol.2014.313

Arroyo-Pardo, E., Baeza, C., Fernández, E., and López-Parra, A. (2007). "Genetic History of the Iberian Peninsula," in *Recent Advances in Molecular Biology and Evolution: Applications to Biological Anthropology*, eds C. Santos and M. Lima (Kerala: Research Signpost), 389–411.

Atzmon, G., Hao, L., Pe'er, I., Velez, C., Pearlman, A., Palamara, P. F., et al. (2010). Abraham's children in the genome era: major Jewish diaspora populations comprise distinct genetic clusters with shared Middle Eastern Ancestry. *Am. J. Hum. Genet.* 86, 850–859. doi: 10.1016/j.ajhg.2010.04.015

Azevedo, E. M. (1994). *Cristãos Novos No Nordeste Transmontano*. Zamora: Asociación Caminos de Cervantes y Sefarad.

Bacon, I. B. A. J. (2011). *The Historical Atlas of Judaism*. London: Chartwell Books.

Balaresque, P., Bowden, G. R., Adams, S. M., Leung, H. Y., King, T. E., Rosser, Z. H., et al. (2010). A predominantly neolithic origin for European paternal lineages. *PLoS Biol.* 8:e1000285. doi: 10.1371/journal.pbio.1000285

Barnavi, E., Eliav-Feldon, M., and Opatowski, M. (1992). *A Historical Atlas of the Jewish People: From the Time of the Patriarchs to the Present*. London: Hutchinson.

Bauchet, M., Mcevoy, B., Pearson, L. N., Quillen, E. E., Sarkisian, T., Hovhannesyan, K., et al. (2007). Measuring European population stratification with microarray genotype data. *Am. J. Hum. Genet.* 80, 948–956. doi: 10.1086/513477

Bedford, F. L. (2011). Sephardic signature in haplogroup T mitochondrial DNA. *Eur. J. Hum. Genet.* 20, 441–448. doi: 10.1038/ejhg.2011.200

Bedford, F., Yacobi, D., Felix, G., and Garza, F. (2013). Clarifying Mitochondrial DNA Subclades of T2e from Mideast to Mexico. *J. Phylogen. Evol. Biol.* 1:2.

Behar, D. M., Garrigan, D., Kaplan, M. E., Mobasher, Z., Rosengarten, D., Karafet, T. M., et al. (2004). Contrasting patterns of Y chromosome variation in Ashkenazi Jewish and host non-Jewish European populations. *Hum. Genet.* 114, 354–365. doi: 10.1007/s00439-003-1073-7

Behar, D. M., Metspalu, E., Kivisild, T., Rosset, S., Tzur, S., Hadid, Y., et al. (2008). Counting the founders: the matrilineal genetic ancestry of the Jewish Diaspora. *PLoS ONE* 3:e2062. doi: 10.1371/journal.pone.0002062

Behar, D. M., Thomas, M. G., Skorecki, K., Hammer, M. F., Bulygina, E., Rosengarten, D., et al. (2003). Multiple origins of Ashkenazi Levites: Y chromosome evidence for both Near Eastern and European ancestries. *Am. J. Hum. Genet.* 73, 768–779. doi: 10.1086/378506

Behar, D. M., Van Oven, M., Rosset, S., Metspalu, M., Loogvali, E. L., Silva, N. M., et al. (2012). A "copernican" reassessment of the human mitochondrial DNA tree from its root. *Am. J. Hum. Genet.* 90, 675–684. doi: 10.1016/j.ajhg.2012.03.002

Behar, D. M., Yunusbayev, B., Metspalu, M., Metspalu, E., Rosset, S., Parik, J., et al. (2010). The genome-wide structure of the Jewish people. *Nature* 466, 238–242. doi: 10.1038/nature09103

Beleza, S., Gusmao, L., Amorim, A., Carracedo, A., and Salas, A. (2005). The genetic legacy of western Bantu migrations. *Hum. Genet.* 117, 366–375. doi: 10.1007/s00439-005-1290-3

Beleza, S., Gusmao, L., Lopes, A., Alves, C., Gomes, I., Giouzeli, M., et al. (2006). Micro-phylogeographic and demographic history of Portuguese male lineages. *Ann. Hum. Genet.* 70, 181–194. doi: 10.1111/j.1529-8817.2005.00221.x

Bonne-Tamir, B., Johnson, M., Natali, A., Wallace, D., and Cavalli-Sforza, L. (1986). Human mitochondrial DNA types in two Israeli populations–a comparative study at the DNA level. *Am. J. Hum. Genet.* 38:341.

Bosch, E., Calafell, F., Comas, D., Oefner, P. J., Underhill, P. A., and Bertranpetit, J. (2001). High-resolution analysis of human Y-chromosome variation shows a sharp discontinuity and limited gene flow between northwestern Africa and the Iberian Peninsula. *Am. J. Hum. Genet.* 68, 1019–1029. doi: 10.1086/319521

Brehm, A., Pereira, L., Kivisild, T., and Amorim, A. (2003). Mitochondrial portraits of the Madeira and Acores archipelagos witness different genetic pools of its settlers. *Hum. Genet.* 114, 77–86. doi: 10.1007/s00439-003-1024-3

Brion, M., Sanchez, J. J., Balogh, K., Thacker, C., Blanco-Verea, A., Borsting, C., et al. (2005). Introduction of an single nucleodite polymorphism-based "Major Y-chromosome haplogroup typing kit" suitable for predicting the geographical origin of male lineages. *Electrophoresis* 26, 4411–4420. doi: 10.1002/elps.200500293

Brown, W. M. (1980). Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. *Proc. Natl. Acad. Sci. U.S.A.* 77, 3605–3609. doi: 10.1073/pnas.77.6.3605

Buades, J., Ben-Chetrit, E., and Levy, M. (1995). Familial Mediterranean fever in the "Chuetas" of Mallorca–origin in inquisition? *Isr. J. Med. Sci.* 31, 497–499.

Busby, G. B., Brisighelli, F., Sanchez-Diz, P., Ramos-Luis, E., Martinez-Cadenas, C., Thomas, M. G., et al. (2011). The peopling of Europe and the cautionary tale of Y chromosome lineage R-M269. *Proc. Biol. Sci.* 279, 884–892. doi: 10.1098/rspb.2011.1044

Cambra, A., Muñoz-Saá, I., Crespí, C., Serra, A., Etxagibel, A., Matamoros, N., et al. (2009). MICA-HLA-B haplotype diversity and linkage disequilibrium in a population of Jewish descent from Majorca (the Balearic Islands). *Hum. Immunol.* 70, 513–517. doi: 10.1016/j.humimm.2009.04.005

Campbell, C. L., Palamara, P. F., Dubrovsky, M., Botigué, L. R., Fellous, M., Atzmon, G., et al. (2012). North African Jewish and non-Jewish populations form distinctive, orthogonal clusters. *Proc. Natl. Acad. Sci. U.S.A.* 109, 13865–13870. doi: 10.1073/pnas.1204840109

Canelo, D. A. (1987). *Os Ultimos Criptojudeus em Portugal*. Belmonte: Centro de Cultura Pedro Alvares Cabral. 21–42.

Cann, R. L., Stoneking, M., and Wilson, A. C. (1987). Mitochondrial DNA and human evolution. *Nature* 325, 31–36. doi: 10.1038/325031a0

Capelli, C., Onofri, V., Brisighelli, F., Boschi, I., Scarnicci, F., Masullo, M., et al. (2009). Moors and Saracens in Europe: estimating the medieval North African male legacy in southern Europe. *Eur. J. Hum. Genet.* 17, 848–852. doi: 10.1038/ejhg.2008.258

Carvajal-Carmona, L. G., Soto, I. D., Pineda, N., Ortiz-Barrientos, D., Duque, C., Ospina-Duque, J., et al. (2000). Strong Amerind/white sex bias and a possible Sephardic contribution among the founders of a population in northwest Colombia. *Am. J. Hum. Genet.* 67, 1287–1295. doi: 10.1086/321216

Carvalho, A. C. (1999). *Os Judeus do desterro de Portugal*. Lisboa: Quetzal, 45–150.
Carvalho-Silva, D. R., Santos, F. R., Rocha, J., and Pena, S. D. (2001). The phylogeography of Brazilian Y-chromosome lineages. *Am. J. Hum. Genet.* 68, 281–286. doi: 10.1086/316931
Cinnioğlu, C., King, R., Kivisild, T., Kalfoğlu, E., Atasoy, S., Cavalleri, G. L., et al. (2004). Excavating Y-chromosome haplotype strata in Anatolia. *Hum. Genet.* 114, 127–148. doi: 10.1007/s00439-003-1031-4
Cruciani, F., Santolamazza, P., Shen, P., Macaulay, V., Moral, P., Olckers, A., et al. (2002). A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. *Am. J. Hum. Genet.* 70, 1197–1214. doi: 10.1086/340257
Di Giacomo, F., Luca, F., Popa, L. O., Akar, N., Anagnou, N., Banyko, J., et al. (2004). Y chromosomal haplogroup J as a signature of the post-neolithic colonization of Europe. *Hum. Genet.* 115, 357–371. doi: 10.1007/s00439-004-1168-9
Domingo, C., Touitou, I., Bayou, A., Ozen, S., Notarnicola, C., Dewalle, M., et al. (2000). Familial Mediterranean fever in the 'Chuetas' of Mallorca: a question of Jewish origin or genetic heterogeneity. *Eur. J. Hum. Genet.* 8, 242–246. doi: 10.1038/sj.ejhg.5200462
Elhaik, E. (2013). The missing link of Jewish European ancestry: Contrasting the Rhineland and the Khazarian hypotheses. *Genome Biol. Evol.* 5, 61–74. doi: 10.1093/gbe/evs119
Elvira, M. (2007). *A Problemática do Judaísmo. Séculos XVI-XVII. Actas do Congresso Internacional Inquisição Portuguesa Tempo, Razão e Circunstância*. Lisboa, 619.
Feldman, G. L., Schrijver, I., Lyon, E., and Palomaki, G. E. (2014). Results of the College of American Pathology/American College of Medical Genetics and Genomics external proficiency testing from 2006 to 2013 for three conditions prevalent in the Ashkenazi Jewish population. *Genet. Med.* 16, 695–702. doi: 10.1038/gim.2014.14
Flores, C., Maca-Meyer, N., Gonzalez, A. M., Oefner, P. J., Shen, P., Perez, J. A., et al. (2004). Reduced genetic structure of the Iberian peninsula revealed by Y-chromosome analysis: implications for population demography. *Eur. J. Hum. Genet.* 12, 855–863. doi: 10.1038/sj.ejhg.5201225
Gerber, S., Rozet, J.-M., Takezawa, S.-I., Dos Santos, L. C., Lopes, L., Gribouval, O., et al. (2000). The photoreceptor cell-specific nuclear receptor gene (PNR) accounts for retinitis pigmentosa in the Crypto-Jews from Portugal (Marranos), survivors from the Spanish Inquisition. *Hum. Genet.* 107, 276–284. doi: 10.1007/s004390000350
Goncalves, R., Freitas, A., Branco, M., Rosa, A., Fernandes, A. T., Zhivotovsky, L. A., et al. (2005). Y-chromosome lineages from Portugal, Madeira and Acores record elements of Sephardim and Berber ancestry. *Ann. Hum. Genet.* 69, 443–454. doi: 10.1111/j.1529-8817.2005.00161.x
Gonzalez, A. M., Brehm, A., Perez, J. A., Maca-Meyer, N., Flores, C., and Cabrera, V. M. (2003). Mitochondrial DNA affinities at the Atlantic fringe of Europe. *Am. J. Phys. Anthropol.* 120, 391–404. doi: 10.1002/ajpa.10168
Guix, P., Picornell, A., Parera, M., Galmes, A., Obrador, A., Ramon, M., et al. (2002). Distribution of HFE C282Y and H63D mutations in the Balearic Islands (NE Spain). *Clin. Genet.* 61, 43–48. doi: 10.1034/j.1399-0004.2002.610109.x
Hammer, M. F. (1994). A recent insertion of an alu element on the Y chromosome is a useful marker for human population studies. *Mol. Biol. Evol.* 11, 749–761.
Hammer, M. F., Behar, D. M., Karafet, T. M., Mendez, F. L., Hallmark, B., Erez, T., et al. (2009). Extended Y chromosome haplotypes resolve multiple and unique lineages of the Jewish priesthood. *Hum. Genet.* 126, 707–717. doi: 10.1007/s00439-009-0727-5
Hammer, M. F., Redd, A. J., Wood, E. T., Bonner, M. R., Jarjanazi, H., Karafet, T., et al. (2000). Jewish and Middle Eastern non-Jewish populations share a common pool of Y-chromosome biallelic haplotypes. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6769–6774. doi: 10.1073/pnas.100115997
Hammer, M. F., Skorecki, K., Selig, S., Blazer, S., Rappaport, B., Bradman, R., et al. (1997). Y chromosomes of Jewish priests. *Nature* 385:32. doi: 10.1038/385032a0
Hernández, C. L., Reales, G., Dugoujon, J.-M., Novelletto, A., Rodríguez, J. N., Cuesta, P., et al. (2014). Human maternal heritage in Andalusia (Spain): its composition reveals high internal complexity and distinctive influences of mtDNA haplogroups U6 and L in the western and eastern side of region. *BMC Genet.* 15:11. doi: 10.1186/1471-2156-15-11
Jobling, M. A., and Tyler-Smith, C. (2003). The human Y chromosome: an evolutionary marker comes of age. *Nat. Rev. Genet.* 4, 598–612. doi: 10.1038/nrg1124
King, T. E., Bowden, G. R., Balaresque, P. L., Adams, S. M., Shanks, M. E., and Jobling, M. A. (2007). Thomas Jefferson's Y chromosome belongs to a rare European lineage. *Am. J. Phys. Anthropol.* 132, 584–589. doi: 10.1002/ajpa.20557
Kopelman, N. M., Stone, L., Wang, C., Gefel, D., Feldman, M. W., Hillel, J., et al. (2009). Genomic microsatellites identify shared Jewish ancestry intermediate between Middle Eastern and European populations. *BMC Genet.* 10:80. doi: 10.1186/1471-2156-10-80
Maca-Meyer, N., Sanchez-Velasco, P., Flores, C., Larruga, J. M., Gonzalez, A. M., Oterino, A., et al. (2003). Y chromosome and mitochondrial DNA characterization of Pasiegos, a human isolate from Cantabria (Spain). *Ann. Hum. Genet.* 67, 329–339. doi: 10.1046/j.1469-1809.2003.00045.x
Mairal, Q., Santos, C., Silva, M., Marques, S. L., Ramos, A., Aluja, M. P., et al. (2013). Linguistic isolates in Portugal: Insights from the mitochondrial DNA pattern. *Forensic Sci. Int. Genet.* 7, 618–623. doi: 10.1016/j.fsigen.2013.08.009
Martins, J. (2006). *Portugal e os Judeus: Judaísmo e Anti-Semitismo no Século XX*. Lisbon: Vega, 115–154.
Maurício, D. (1971). "Judeus," in *Enciclopédia Luso-Brasileira de Cultura* (Lisboa: Verbo), 806–842.
Mendizabal, I., Valente, C., Gusmão, A., Alves, C., Gomes, V., Goios, A., et al. (2011). Reconstructing the Indian origin and dispersal of the European Roma: a maternal genetic perspective. *PLoS ONE* 6:e15988. doi: 10.1371/journal.pone.0015988
Moore, L. T., Mcevoy, B., Cape, E., Simms, K., and Bradley, D. G. (2006). A Y-chromosome signature of hegemony in Gaelic Ireland. *Am. J. Hum. Genet.* 78, 334–338. doi: 10.1086/500055
Nebel, A., Filon, D., Brinkmann, B., Majumder, P. P., Faerman, M., and Oppenheim, A. (2001). The Y chromosome pool of Jews as part of the genetic landscape of the Middle East. *Am. J. Hum. Genet.* 69, 1095–1112. doi: 10.1086/324070
Nebel, A., Filon, D., Weiss, D. A., Weale, M., Faerman, M., Oppenheim, A., et al. (2000). High-resolution Y chromosome haplotypes of Israeli and Palestinian Arabs reveal geographic substructure and substantial overlap with haplotypes of Jews. *Hum. Genet.* 107, 630–641. doi: 10.1007/s004390000426
Nogueiro, I., Manco, L., Gomes, V., Amorim, A., and Gusmao, L. (2010). Phylogeographic analysis of paternal lineages in NE Portuguese Jewish communities. *Am. J. Phys. Anthropol.* 141, 373–381. doi: 10.1002/ajpa.21154
Nogueiro, I., Teixeira, J., Amorim, A., Gusmao, L., and Alvarez, L. (2014). Echoes from Sepharad: signatures on the maternal gene pool of crypto-Jewish descendants. *Eur. J. Hum. Genet.* doi: 10.1038/ejhg.2014.140 [Epub ahead of print].
Oefner, P. J., Holzi, G., Shen, P., Shpirer, I., Gefel, D., Lavi, T., et al. (2013). Genetics and the history of the Samaritans: Y-chromosomal microsatellites and genetic affinity between Samaritans and Cohanim. *Hum. Biol.* 85, 825–858. doi: 10.3378/027.085.0601
Olshen, A. B., Gold, B., Lohmueller, K. E., Struewing, J. P., Satagopan, J., Stefanov, S. A., et al. (2008). Analysis of genetic variation in Ashkenazi Jews by high density SNP genotyping. *BMC Genet.* 9:14. doi: 10.1186/1471-2156-9-14
Ostrer, H. (2001). A genetic profile of contemporary Jewish populations. *Nat. Rev. Genet.* 2, 891–898. doi: 10.1038/35098506
Ostrer, H., and Skorecki, K. (2013). The population genetics of the Jewish people. *Hum. Genet.* 132, 119–127. doi: 10.1007/s00439-012-1235-6
Pacheco, P. R., Branco, C. C., Cabral, R., Costa, S., Araujo, A. L., Peixoto, B. R., et al. (2005). The Y-chromosomal heritage of the Azores Islands population. *Ann. Hum. Genet.* 69, 145–156. doi: 10.1046/j.1469-1809.2004.00147.x
Pala, M., Olivieri, A., Achilli, A., Accetturo, M., Metspalu, E., Reidla, M., et al. (2012). Mitochondrial DNA signals of late glacial recolonization of Europe from near eastern refugia. *Am. J. Hum. Genet.* 90, 915–924. doi: 10.1016/j.ajhg.2012.04.003
Paulo, A. (1985). *Os Judeus Secretos em Portugal*. Porto: Editorial Labirinto, 5–19.
Pereira, L., Prata, M. J., and Amorim, A. (2000). Diversity of mtDNA lineages in Portugal: not a genetic edge of European variation. *Ann. Hum. Genet.* 64, 491–506. doi: 10.1046/j.1469-1809.2000.6460491.x
Pereira, V., Gomes, V., Amorim, A., Gusmão, L., and João Prata, M. (2010). Genetic characterization of uniparental lineages in populations from Southwest Iberia with past malaria endemicity. *Am. J. Hum. Biol.* 22, 588–595. doi: 10.1002/ajhb.21049
Picornell, A., Gimenez, P., Castro, J. A., and Ramon, M. M. (2006). Mitochondrial DNA sequence variation in Jewish populations. *Int. J. Legal Med.* 120, 271–281. doi: 10.1007/s00414-006-0083-0

Picornell, A., Gomez-Barbeito, L., Tomas, C., Castro, J. A., and Ramon, M. M. (2005). Mitochondrial DNA HVRI variation in Balearic populations. *Am. J. Phys. Anthropol.* 128, 119–130. doi: 10.1002/ajpa.10423

Pignatelli, M. (2000). *A Comunidade Israelita de Lisboa : o Passado e o Presente na Construção da Etnicidade dos Judeus de Lisboa*. Lisboa: Universidade Técnica de Lisboa, 33–89.

Rosner, G., Rosner, S., and Orr-Urtreger, A. (2009). Genetic testing in Israel: an overview. *Annu. Rev. Genomics Hum. Genet.* 10, 175–192. doi: 10.1146/annurev.genom.030308.111406

Roth, C. (1979). *Historia de los Marranos*. (Madrid: Altalena), 163–184.

Santamaria Arández, Á. (1997). "Sobre la condición de los conversos y chuetas de Mallorca," in *Serie III. Historia Medieval 10* (Mallorca: Espacio, Tiempo y Forma), 219–262.

Santos, C., Fregel, R., Cabrera, V. M., Alvarez, L., Larruga, J. M., Ramos, A., et al. (2014). Mitochondrial DNA and Y-chromosome structure at the Mediterranean and Atlantic facades of the Iberian Peninsula. *Am. J. Hum. Biol.* 26, 130–141. doi: 10.1002/ajhb.22497

Santos, C., Fregel, R., Cabrera, V. M., Gonzalez, A. M., Larruga, J. M., and Lima, M. (2010). Mitochondrial DNA patterns in the Macaronesia islands: variation within and among archipelagos. *Am. J. Phys. Anthropol.* 141, 610–619. doi: 10.1002/ajpa.21180

Santos, C., Lima, M., Montiel, R., Angles, N., Pires, L., Abade, A., et al. (2003). Genetic structure and origin of peopling in the Azores islands (Portugal): the view from mtDNA. *Ann. Hum. Genet.* 67, 433–456. doi: 10.1046/j.1469-1809.2003.00031.x

Seldin, M. F., Shigeta, R., Villoslada, P., Selmi, C., Tuomilehto, J., Silva, G., et al. (2006). European population substructure: clustering of northern and southern populations. *PLoS Genet.* 2:e143. doi: 10.1371/journal.pgen.0020143

Semino, O., Magri, C., Benuzzi, G., Lin, A. A., Al-Zahery, N., Battaglia, V., et al. (2004). Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the Mediterranean area. *Am. J. Hum. Genet.* 74, 1023–1034. doi: 10.1086/386295

Semino, O., Passarino, G., Oefner, P. J., Lin, A. A., Arbuzova, S., Beckman, L. E., et al. (2000). The genetic legacy of Paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science* 290, 1155–1159. doi: 10.1126/science.290.5494.1155

Shen, P., Lavi, T., Kivisild, T., Chou, V., Sengun, D., Gefel, D., et al. (2004). Reconstruction of patrilineages and matrilineages of Samaritans and other Israeli populations from Y-chromosome and mitochondrial DNA sequence variation. *Hum. Mutat.* 24, 248–260. doi: 10.1002/humu.20077

Sutton, W. K., Knight, A., Underhill, P. A., Neulander, J. S., Disotell, T. R., and Mountain, J. L. (2006). Toward resolution of the debate regarding purported crypto-Jews in a Spanish-American population: evidence from the Y chromosome. *Ann. Hum. Biol.* 33, 100–111. doi: 10.1080/03014460500475870

Tafe, L. J., Datto, M. B., Palomaki, G. E., and Lacbawan, F. L. (2015). Molecular testing for the *BRCA1* and *BRCA2* Ashkenazi Jewish founder mutations: a report on the College of American Pathologists proficiency testing surveys. *Genet. Med.* 17, 58–62. doi: 10.1038/gim.2014.77

Tavares, M. J. F. (2004). A Herança Judaica em Portugal. Lisboa: CTT Corrreios de Portugal, 79–91.

Teixeira, J. C., Nogueiro, I., Goios, A., Gusmão, L., Amorim, A., and Alvarez, L. (2011). Mitochondrial DNA-control region sequence variation in the NE Portuguese Jewish community. *Forensic Sci. Int. Genet. Suppl. Ser.* 3, e51–e52. doi: 10.1016/j.fsigss.2011.08.025

Thomas, M. G., Parfitt, T., Weiss, D. A., Skorecki, K., Wilson, J. F., Le Roux, M., et al. (2000). Y chromosomes traveling south: the cohen modal haplotype and the origins of the Lemba—the "Black Jews of Southern Africa". *Am. J. Hum. Genet.* 66, 674–686. doi: 10.1086/302749

Thomas, M. G., Skorecki, K., Ben-Ami, H., Parfitt, T., Bradman, N., and Goldstein, D. B. (1998). Origins of Old Testament priests. *Nature* 394, 138–140. doi: 10.1038/28083

Thomas, M. G., Weale, M. E., Jones, A. L., Richards, M., Smith, A., Redhead, N., et al. (2002). Founding mothers of Jewish communities: geographically separated Jewish groups were independently founded by very few female ancestors. *Am. J. Hum. Genet.* 70, 1411–1420. doi: 10.1086/340609

Tikochinski, Y., Ritte, U., Gross, S., Prager, E., and Wilson, A. (1991). mtDNA polymorphism in two communities of Jews. *Am. J. Hum. Genet.* 48:129.

Tofanelli, S., Taglioli, L., Bertoncini, S., Francalacci, P., Klyosov, A., and Pagani, L. (2014). Mitochondrial and Y chromosome haplotype motifs as diagnostic markers of Jewish ancestry: a reconsideration. *Front. Genet.* 5:384. doi: 10.3389/fgene.2014.00384

Underhill, P. A., Passarino, G., Lin, A. A., Shen, P., Mirazon Lahr, M., Foley, R. A., et al. (2001). The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann. Hum. Genet.* 65, 43–62. doi: 10.1046/j.1469-1809.2001.6510043.x

van Oven, M., and Kayser, M. (2009). Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.* 30, E386–E394. doi: 10.1002/humu.20921

Velez, C., Palamara, P. F., Guevara-Aguirre, J., Hao, L., Karafet, T., Guevara-Aguirre, M., et al. (2012). The impact of Converso Jews on the genomes of modern Latin Americans. *Hum. Genet.* 131, 251–263. doi: 10.1007/s00439-011-1072-z

Venturini, G., Koskiniemi-Kuendig, H., Harper, S., Berson, E. L., and Rivolta, C. (2014). Two specific mutations are prevalent causes of recessive retinitis pigmentosa in North American patients of Jewish ancestry. *Genet Med.* doi: 10.1038/gim.2014.132 [Epub ahead of print].

Wilke, C. L. (2009). *História dos Judeus Portugueses*. Lisboa: EDIÇÕES 70, 11–70.

Zalloua, P. A., Platt, D. E., El Sibai, M., Khalife, J., Makhoul, N., Haber, M., et al. (2008). Identifying genetic traces of historical expansions: Phoenician footprints in the Mediterranean. *Am. J. Hum. Genet.* 83, 633–642. doi: 10.1016/j.ajhg.2008.10.012

Zelinger, L., Banin, E., Obolensky, A., Mizrahi-Meissonnier, L., Beryozkin, A., Bandah-Rozenfeld, D., et al. (2011). A missense mutation in DHDDS, encoding dehydrodolichyl diphosphate synthase, is associated with autosomal-recessive retinitis pigmentosa in Ashkenazi Jews. *Am. J. Hum. Genet.* 88, 207–215. doi: 10.1016/j.ajhg.2011.01.002

Zuchner, S., Dallman, J., Wen, R., Beecham, G., Naj, A., Farooq, A., et al. (2011). Whole-exome sequencing links a variant in DHDDS to retinitis pigmentosa. *Am. J. Hum. Genet.* 88, 201–206. doi: 10.1016/j.ajhg.2011.01.001

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 18 October 2014; accepted: 11 January 2015; published online: 02 February 2015.

*Citation: Nogueiro I, Teixeira JC, Amorim A, Gusmão L and Alvarez L (2015) Portuguese crypto-Jews: the genetic heritage of a complex history. Front. Genet. **6**:12. doi: 10.3389/fgene.2015.00012*

This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics.

Extended Abstract:

Nogueiro, I., Neto, C., Marques, S. L., Alves, C., Cohen-Addad, N., Amorim, A., and Alvarez, L. (2015). Exploring Sephardic lineages in São Tomé e Príncipe. *Forensic Science International: Genetics Supplement Series*.

G Model
FSIGSS 1179 No. of Pages 3
ARTICLE IN PRESS
Forensic Science International: Genetics Supplement Series xxx (2015) xxx–xxx
Contents lists available at ScienceDirect
Forensic Science International: Genetics Supplement Series
journal homepage: www.elsevier.com/locate/FSIGSS

Exploring Sephardic lineages in São Tomé e Príncipe

Inês Nogueiro[a,b,c,*], Célia Neto[a,b], Sofia L. Marques[a,b], Cíntia Alves[a,b], Nicole Cohen-Addad, António Amorim[a,b,c], Leonor Gusmão[a,b,d], Luis Alvarez[a,b]

[a] Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal
[b] Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP), Porto, Portugal
[c] Faculty of Sciences, University of Porto, Porto, 4169-007, Portugal
[d] DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Rio de Janeiro 21550-013, Brazil

ARTICLE INFO

Article history:
Received 27 August 2015
Accepted 20 September 2015
Available online xxx

Keywords:
Jewish ancestry
Y chromosome
mtDNA
AIMs
São Tomé

ABSTRACT

São Tomé and Príncipe are the two main islands of a small archipelago located in the Gulf of Guinea, western equatorial African coast. These islands were probably uninhabited at the time of the Portuguese discovery in 1471. After the Portuguese decree of expulsion, many new-Christians fled to São Tomé, since the inquisition was never established there. Several documents attest the continuous movement of new-Christians to these islands, which worked as a refuge from inquisitional prosecutions. To elucidate the genetic impact of the historical Jewish migrations to São Tomé, we selected a sample of unrelated individuals based on the following criteria: (a) sharing surnames with those included in the historical reports as Jewish migrants (b) showing cultural practices putatively related to the Jewish religion. Both maternal and paternal lineages were investigated using the complete mtDNA control region, along with 22 Y-SNPs and 23 Y-STR markers. Moreover, 83 autosomal insertion/deletion markers were analyzed.

1. Introduction

São Tomé islands were probably uninhabited at the time of the Portuguese discovery in 1471. Soon after that, the first settlements took place in 1486, by João de Paiva, who took with him many Jews. One of the most controversial episodes of this colonization is the forced migration of Jewish children in 1492 [1]. The information about this occurrence is scarce and diverges between Jewish and Christian historians. Around 2000 young children and teenagers were taken from their families, forcedly baptized [1] and sent to the archipelago so they could be away from the Jewish faith, grow up with a Christian education and contribute to the population of the new territory, one of the main aims of the Portuguese king Joao III [1]. Several documents confirm that many of these children survived and along with other settlers, mainly of African origin, started the process of miscegenation. Over more than a century several documents attest to the continuous movement of new-Christians to São Tomé, which worked as a refuge from the inquisitional prosecutions and also as a strategical location for the flourish intercontinental trade lead manly by New Christians [2].

Despite the documents above mentioned, the information available on the impact or number of Jewish people that really went (and when) to São Tomé is scarce.

In previous genetic studies on lineage markers [3–6] related to the colonization processes, the European impact at the mitochondrial pool was virtually nil, contrasting to approximately 8–24% of Y chromosome lineages in São Tomé. Among the European Y chromosomes, none were found to belong to haplogroups J and T previously reported to have high frequencies in Portuguese Jews (>50%) [7]. In summary, in previous studies, neither maternal nor paternal Jewish founder lineages were detected in the general population from São Tomé. If there is any genetic Jewish heritage still present in São Tomé, different sampling strategies and more genetic markers would be therefore required to trace it.

2. Materials and methods

Population samples from presumed Jewish descendants in São Tomé were selected using a combination of geographic, ethno-historical, and genealogical criteria. Samples from buccal cells were collected on cytology brushes from 22 unrelated individuals, under informed consent. DNA was extracted using a standard Chelex-100 (Bio-Rad, Hercules, CA, USA) procedure.

Twenty two Y-SNP [8–11] and 23 Y-STR markers included in Powerplex Y23 System (Promega Corporation, Madison, WI, USA)

* Corresponding author at: Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP), Porto, Portugal.
E-mail address: inogueiro@ipatimup.pt (I. Nogueiro).

http://dx.doi.org/10.1016/j.fsigss.2015.09.182
1875-1768/

Please cite this article in press as: I. Nogueiro, et al., Exploring Sephardic lineages in São Tomé e Príncipe, Forensic Sci. Int. Gene. Suppl. (2015), http://dx.doi.org/10.1016/j.fsigss.2015.09.182

G Model
FSIGSS 1179 No. of Pages 3
ARTICLE IN PRESS

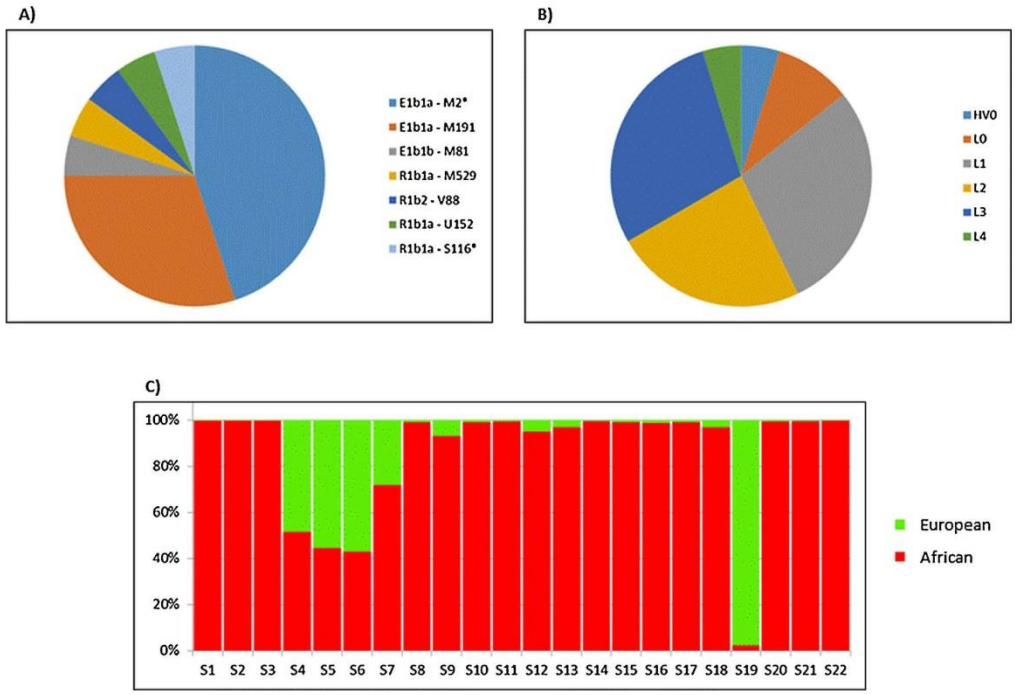


Fig. 1. Haplogroup composition obtained for the putative Jewish descendants from São Tomé: mitochondrial lineages (A); Y chromosome lineages (B), as well as proportions for the African and European clusters within São Tomé population (C) obtained for a total of 83 non-coding autosomal bi-allelic Indels.

were typed. The Y-SNP haplogroups were designated according to van Oven et al. [12], and the Y-STR alleles, according to the ISFG recommendations [13].

A 3348 bp fragment containing the entire mtDNA-control region was also surveyed [14]. All samples were typed for a total number of 83 non-coding autosomal bi-allelic Indels [15,16].

Haplogroup and haplotype frequencies were calculated through direct counting. For autosomal Indels, cluster analysis were performed with STRUCTURE [17], assuming a contribution from Europeans and Africans (i.e., $K = 2$); and 246 individuals from HGDP–CEPH were used as reference. Runs consisted of 100,000 burnin steps and 100,000 MCMC iterations, using the option "Use population Information to test for migrants" with the Admixture model.

3. Results and discussion

Most mtDNA haplogroups detected in our sample are of African origin (95%); a single sample belongs to the European haplogroup HV0 (Fig. 1A) which is not included in the HV0b clade of the Portuguese Jewish founding lineage [18]. Although higher than that of the mtDNA, the frequency of European Y chromosomes (20%) was not significantly different from that previously observed in samples from the general population [6,8] (Fig. 1B). Moreover, the putative Jewish haplogroups J and T were not detected either.

The proportions of African and European ancestries obtained in the analysis of autosomal Indels were of 86% and 14%, respectively (Fig. 1C). The European ancestry was just slightly higher than the expected based on the average value calculated for the mtDNA and Y chromosome haplogroups (12.5%). This can be the result of diverse sex biased admixture events taking place in different periods after the colonization of São Tomé.

Finally, a significantly higher European ancestry was found for the two individuals carrying the European Y chromosomes R1b1-M529 and R1b1-U152, when compared to the average population. This supports a recent admixture rather than an old European influx to the archipelago.

4. Conclusion

In the present study, using a large set of both recombining and non-recombining markers, no significant differences were detected between the alleged Jewish descendants and the host population, showing that no particular database is required in forensic cases in relation to Jewish ancestry, contrasting with previous findings on population heterogeneity when other ethnic criteria were considered.

Conflict of interest

None.

Role of funding

This work was supported by FCT, financed by the European Social Funds (COMPETE FEDER), and project PTDC/ATP DEM/4545/2012. LA (SFRH/BPD/65000/2009), IN (SFRH/BD/73336/2010) supported by FCT fellowships and funded by POPH-QREN, (http://www.poph.qren.pt/).

Please cite this article in press as: I. Nogueiro, et al., Exploring Sephardic lineages in São Tomé e Príncipe, Forensic Sci. Int. Gene. Suppl. (2015), http://dx.doi.org/10.1016/j.fsigss.2015.09.182

G Model
FSIGSS 1179 No. of Pages 3
ARTICLE IN PRESS
I. Nogueiro et al. / Forensic Science International: Genetics Supplement Series xxx (2015) xxx–xxx e3

Acknowledgment

We thank all DNA donors for their collaboration.

References

[1] E. Mea, O resgate dos meninos de S. Tomé em Orion, Literatura e História—Actas do Colóquio Internacional, Porto, 2004, pp. 25–39.
[2] Y. Kaplan, Dicionário do Judaismo Português, Presença, Lisboa, 2009, pp. 174–176 192–206.
[3] M.J. Trovoada, et al., Pattern of mtDNA variation in three populations from Sao Tome e Principe, Ann. Hum. Genet. 68 (Pt. (1)) (2004) 40–54.
[4] M.J. Trovoada, et al., Dissecting the genetic history of Sao Tome e Principe: a new window from Y-chromosome biallelic markers, Ann. Hum. Genet. 71 (Pt. (1)) (2007) 77–85.
[5] M.J. Trovoada, et al., Evidence for population sub-structuring in Sao Tome e Principe as inferred from Y-chromosome STR analysis, Ann. Hum. Genet 65 (Pt. (3)) (2001) 271–283.
[6] R. Gonçalves, H. Spinola, A. Brehm, Y-chromosome lineages in São Tomé e Príncipe islands: evidence of European influence, Am. J. Hum. Biol. 19 (3) (2007) 422–428.
[7] I. Nogueiro, et al., Phylogeographic analysis of paternal lineages in NE Portuguese Jewish communities, Am. J. Phys. Anthropol. 141 (3) (2010) 373–381.
[8] M. Brion, Y chromosome SNP analysis using the single-base extension: a hierarchical multiplex design, Methods Mol. Biol. 297 (2005) 229–242.
[9] V. Gomes, et al., Digging deeper into East African human Y chromosome lineages, Hum. Genet. 127 (5) (2010) 603–613.
[10] A. Gusmao, et al., A Perspective on the history of the Iberian Gypsies provided by phylogeographic analysis of Y-chromosome lineages, Ann. Hum. Genet. 72 (2) (2008) 215–227.
[11] M. González, et al., The genetic landscape of Equatorial Guinea and the origin and migration routes of the Y chromosome haplogroup R-V88, Eur. J. Hum. Genet. 21 (3) (2013) 324–331.
[12] M. Oven, et al., Seeing the wood for the trees: a minimal reference phylogeny for the human Y chromosome, Hum. Mutat. 35 (2) (2014) 187–191.
[13] L. Gusmao, et al., DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis, Int. J. Legal Med. 120 (4) (2006) 191–200.
[14] A. Ramos, et al., Human mitochondrial DNA complete amplification and sequencing: a new validated primer set that prevents nuclear DNA sequences of mitochondrial origin co-amplification, Electrophoresis 30 (9) (2009) 1587–1593.
[15] R. Pereira, et al., Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing, PLoS One 7 (1) (2012) e29684.
[16] R. Pereira, et al., A new multiplex for human identification using insertion/deletion polymorphisms, Electrophoresis 30 (21) (2009) 3682–3690.
[17] M.J. Hubisz, et al., Inferring weak population structure with the assistance of sample group information, Mol. Ecol. Resour. 9 (5) (2009) 1322–1332.
[18] I. Nogueiro, et al., Echoes from Sepharad: signatures on the maternal gene pool of crypto-Jewish descendants, Eur. J. Hum. Genet. (2014) .

Please cite this article in press as: I. Nogueiro, et al., Exploring Sephardic lineages in São Tomé e Príncipe, Forensic Sci. Int. Gene. Suppl. (2015), http://dx.doi.org/10.1016/j.fsigss.2015.09.182

Research article:

Inês Nogueiro, Célia Neto, Sofia L. Marques, António Amorim, Leonor Gusmão, Luis Alvarez (2015). Unveiling the Sephardic Portuguese genetic legacy from autosomal and X chromosome STRs markers (Submitted)

Unveiling the Sephardic Portuguese genetic legacy from autosomal and X-chromosomal STR markers

Inês Nogueiro[1,2,3]*, Célia Neto[1,3], Sofia L. Marques[1,3], António Amorim[1,2,3], Leonor Gusmão[1,3,4] Luis Alvarez[1,3]

[1]Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP), Porto, 4200-465, Portugal

[2]Faculty of Sciences, University of Porto, Porto, 4169-007, Portugal

[3] Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal

[4]DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Rio de Janeiro, 21550-013, Brazil

Number of text pages, plus bibliography: 15

Number of figures: 5

Number of tables: 5

Abbreviated title: **Recombining markers in Portuguese Jews**

*** Correspondence:** Inês Nogueiro, Ipatimup, Rua Dr. Roberto Frias, s/n, 4200-465 Porto, PORTUGAL. Tel.: +351225570700 Fax : +351225570799
inogueiro@ipatimup.pt

Keywords: Crypto-Jews, AmpFLSTR(®) Identifiler™, Argus X-12

This work was supported by FCT, Portuguese Foundation for Science and Technology, financed by the European Social Funds (COMPETE-FEDER), and project PTDC/ATP-DEM/4545/2012. LA (SFRH/BPD/65000/2009), IN (SFRH/BD/73336/2010) are supported by FCT fellowships and funded by POPH-QREN – Promotion of scientific employment, supported by the European Social Fund and national funds of the Ministry of Education and Science. IPATIMUP is an Associate Laboratory of the Portuguese Ministry of Science, Technology and Higher Education and integrates the i3S Research Unit, which is partially supported by FCT,

Abstract

Objectives:

Genetic studies on Sephardim, the heterogeneous group of Jews with Iberian origin, are very scarce. Previous studies on uniparental markers showed a remarkable genetic diversity, a significant Sephardic signature, as well as signs of admixture among the Portuguese NE Jewish communities; here we deepen and refine their genetic portrait using recombining (autosomal and X-linked) markers.

Materials and Methods:

A sample of 54 Portuguese Jews was analyzed using a set of microsatellite (STR) markers (15 autosomal and 12 on the X chromosome). For comparative purposes, 87 individuals from a linguistic isolate on the same geographical area as well as other Portuguese non-Jews were newly typed for autosomal STRs making up a total of 245 individuals.

Results:

High genetic diversity levels were found both for autosomal as well as for the X chromosome, despite a tendency for the loss of alleles. Significant linkage disequilibrium was detected in four pairs of loci for the Portuguese Jews. A sex biased effective population size over time in the Jewish population was also detected.

Discussion:

As for lineage markers, the Portuguese Jewish communities display high levels of genetic diversity at the recombining genome; the diversity decline expected due to genetic drift (evidenced by the reduction of the number of alleles per locus) is mitigated by admixture with the non-Jews. Moreover, the simultaneous analysis of both autosomal and heterosomal markers revealed a sex biased demographic history, suggesting an asymmetry between female and male effective population sizes in the admixture process between Jews and non-Jews.

Introduction

Genetic diversity patterns in modern human populations are the outcome of different demographic events as migrations, colonizations and expansions, and also of evolutionary forces, as mutation, genetic drift and selection (Balaresque et al. 2007; Jobling 2012; Zegura 2005). The reconstruction of human demographic history is based on different kinds of genetic systems, namely autosomes, X chromosome, non-recombining portion of the Y chromosome (NRY) and mitochondrial DNA (mtDNA). Each of these genetic systems have particular features such as differences in copy number and inheritance patterns, as well as in recombination and mutation rates (Garrigan and Hammer 2006).

Assessment of genetic diversity of both the X chromosome and autosomes can contribute to disclose the possible differences between male and female demographic histories and the factors that have shaped their genetic profiles in a population (Arbiza et al. 2014; Emery et al. 2010; Hammer et al. 2008; Heyer et al. 2012; Jobling 2012). In males, there is only one copy of the X chromosome while females possess two, whereas autosomes exist in two copies in both males and females. Assuming that both genders contribute equally to the next generation, the effective population size of the X chromosome is 3/4 that of the autosomes. With the exception of pseudoautosomal regions, the X chromosome only recombines in females, and consequently the recombination rate is lower than that of the autosomes and linkage disequilibrium (LD) tends to be greater (Garrigan and Hammer 2006; Schaffner 2004).

Concerning population genetic studies at a micro-geographical scale, the use of uniparental markers has been far more common than the use of X chromosome or autosomal data. Since both NRY and mtDNA have well established geographical distribution patterns and phylogenies (Jobling and Tyler-Smith 2003; Karafet et al. 2008; Underhill and Kivisild 2007; van Oven and Kayser 2009), they can decode distinct genealogical histories, male and female respectively. Nonetheless, haploid systems are more susceptible to genetic drift, which may obscure historical reconstructions, and produce distinct histories for male and female lineages of the same population, even in the absence of a gender-biased demography. Autosomes and the X chromosome do have recombination that reshuffles lineages in each generation, and if this could be taken as an additional difficulty, on the other hand, it also provides tools to

the reconstruction of human evolution history delivering information unavailable from uniparental markers.

In previous studies we assessed Y chromosome (Nogueiro et al. 2010) and mtDNA diversity patterns (Nogueiro et al. 2014; Teixeira et al. 2011) of the Portuguese crypto Jewish descendants. Crypto-Judaism, (the secret adherence to Judaism while publicly professing another faith), arose in Portugal in the beginning of the 16th century after the Decree of Expulsion and the establishment of the Inquisition. Surprisingly, the scientific community acknowledged the persistence of Crypto-Judaic communities only at the beginning of the 20th century in central and northeastern regions of Portugal, namely in the Bragança district [for a review, see Nogueiro et al. (2015)]. This phenomenon of crypto-Judaism, besides the obvious cultural and anthropological interest, represents from the population genetics point of view, a treasurable natural experiment on how the persistence of hidden religious practices and a strong sense of belonging to a community can be translated into genetics.

Previous results from both uniparental markers showed an unexpected level of genetic diversity for such a small and isolated community, when compared to other Jewish groups, particularly to another Portuguese Sephardic community, Belmonte (Adams et al. 2008; Behar et al. 2008). The term "Sephardic" refers to the Jewish communities that can trace back their origin in Iberia (*Sepharad* in Hebrew), but it can also include a broader definition that comprises the Jewish communities established along the Mediterranean and Middle East.

To complement what the first studies reveled for uniparental markers, in the present study we extended the characterization of the same sample of Portuguese Jews to a set of 15 autosomal and 12 X chromosome Short Tandem Repeat markers (STRs).

Materials and Methods

Samples

Fifty four unrelated self-designated Jewish male individuals from the Bragança district, previously typed for the NRY (Nogueiro et al. 2010) and mtDNA (Nogueiro et al. 2014; Teixeira et al. 2011), were studied. Sampling criteria and collection method, as well as DNA extraction, are described in Nogueiro et al. (2010). The study was approved by the Ethics Committee of the University of Porto (Nº02/CEUP/2012) and appropriate informed consent was required from all subjects.

For comparative purposes, patterns of genetic diversity for the autosomal markers were also newly addressed in 87 individuals of a linguistic isolated population from the same geographical area (Miranda do Douro), and as a reference non-Jewish host population we used the sampling from a previous work (Amorim et al. 2006) as well as newly typed samples from northern Portugal, making up a total of 245 individuals.

Besides the non-Jewish Portuguese population, autosomal data comparisons were also made with other Jewish populations: Morocco, Tunisia, Libya, Iran, Iraq, Turkey and the Chueta Jewish community of Mallorca (Picornell et al. 2003; Picornell et al. 2002), using a set of 9 common loci (D3S1358, D5S818, D7S820, D8S1179, D13S317, D18S51, D21S11, FGA and vWA). In this study the broader concept of "Sephardic" will be used to designate the populations analyzed here (including the oriental Jews from Iran and Iraq). Corresponding non-Jewish populations (NJ) from Morocco (Coudray et al. 2007), Tunisia (Khodjet-El-Khil et al. 2008), Libya (Khodjet-el-Khil et al. 2012), Iran (Shepard and Herrera 2006), Iraq (Barni et al. 2007), Turkey (Bulbul et al. 2014) and the Balearic Islands (Mallorca) (Tomàs et al. 2000) were also included.

Since no data were available concerning the X chromosome markers for other Jewish groups until very recently (Ferragut et al. 2015), population comparisons were made more exhaustively with the linguistic isolates from Miranda do Douro and a Spanish geographical neighbor population from Zamora (Pinto et al. 2014) as well as a reference general Portuguese population (Gomes et al. 2012). The geographical location of each population is presented in Supplementary Figure S1.

DNA typing and analyses

Fifteen autosomal STRs loci (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, vWA, TPOX, D18S51, D5S818, FGA, D2S1338, and D19S433) were typed using the AmpFlSTR®Identifiler®PCR amplification kit (AB Applied Biosystems, Foster City, CA). Twelve X chromosome STRs loci (DXS10148, DXS10135, DXS8378, DXS7132, DXS10079, DXS10074, DXS10103, HPRTB, DXS10101, DXS10146, DXS10134, and DXS7423) included in the Investigator Argus X-12 kit (Qiagen, Hilden) were also typed following the manufacturer's instructions. The amplified products were separated in an ABI 3100 Genetic Analyzer (Thermo Fisher Scientific), and genotyping was performed using GeneMapper 4.0 software (Thermo Fisher Scientific).

Since only male individuals were examined in this study, for X-chromosomal markers haplotypes were directly counted, and no tests for Hardy–Weinberg equilibrium were made. Allele frequencies, mean numbers of pairwise differences and expected gene diversities for autosomal and X-chromosomal markers were assessed using Arlequin v.3.5 software (Excoffier and Lischer 2010). Deviations from Hardy–Weinberg equilibrium for autosomal markers were also addressed with Arlequin v.3.5 software (Excoffier and Lischer 2010). Moreover, pairwise LD analysis between pairs of loci were tested with Arlequin v.3.5 software (Excoffier and Lischer 2010), using 10^6 steps in the Markov chain analysis.

To further investigate the observed genetic diversities, allelic variation per loci was also estimated. For this purpose, the effect of the different population sample sizes was mitigated through bootstrap resampling procedures implemented in R (Team 2014).

Pairwise genetic distances (F_{ST}) were calculated for the Portuguese Jews, Miranda do Douro and Northern Portuguese population using genotypic autosomal data with Arlequin v.3.5 software. F_{ST} were also calculated for the X-chromosomal markers for the Bragança Jews and the reference non-Jewish populations (Miranda do Douro, Zamora and Portugal).

Furthermore, genetic diversities and pairwise F_{ST} values were calculated based on allele frequencies including the 9 common autosomal STR markers using Poptree2 software (Takezaki et al. 2010) for all the Jewish populations considered here and their respective non-Jewish host populations, since only frequency data were available for all populations. The obtained distance matrix was represented in a multidimensional scaling (MDS) using SPSS ver. 21.0 (IBM Corp. Released 2012). Correction for multiple tests was performed according to Bonferroni (1936).

Results

Autosomal STRs

Observed allele frequencies at the analyzed 15 autosomal STR loci are summarized in Table S1. The observed and expected heterozygosity, and Hardy-Weinberg equilibrium (HW) P values for the Portuguese Jews, Miranda do Douro and Northern Portuguese populations are presented in Table 1. For the Bragança Jewish population, no significant

deviations from Hardy-Weinberg equilibrium were observed, after Bonferroni's correction (15 independent tests, significance value of 0.0033) except for locus D19S433 ($P = 0.0032$).

For the Jewish group, with the exception of D5S818, D7S820 and FGA loci, the overall genotype distributions showed an excess of homozygotes compared to the reference host population.

Gene diversities calculated for all loci (Table 1) were compared with the Portuguese non-Jewish population of Miranda do Douro as well as those reported for the northern Portugal. In the Portuguese Jews, values ranged from 69% for TPOX to 88% for D18S51, with an average of 79.3% for all 15 STRs. For the same set of markers, the average diversity was 79.0% in Miranda do Douro and 79.5% in the non-Jewish northern Portuguese.

Genetic diversities of a set of 9 shared STR markers were also compared between the Portuguese Jews and its host populations, and other Sephardic Jews and their respective non-Jewish host populations (data not shown). Considering the Jewish populations, values ranged from 74% for the Libyan Jews, 81% for Tunisia and the Jewish Chueta community, 82% in the Bragança Jews, Iraq, Morocco and Turkey, to the most diverse Jewish community of Iran with 85%. Respective non-Jewish (NJ) host populations showed a slightly shorter range of diversities, varying from 80% in Morocco and Tunisia, 81% in Miranda do Douro, Iraq, Turkey, Libya and Iran to a maximum of 82% in the Mallorca Islands and northern Portugal.

Overall the genetic diversities between the different Jewish communities and the respective non-Jewish host populations do not reveal major differences. To further investigate this issue, the number of allele losses in each of the nine loci was addressed in all populations. Figure 1 shows the results obtained for the Portuguese Jews, Miranda do Douro and Northern Portugal. Results for all other populations can be seen in Supplementary Figure S2.

Pairwise genetic distances (F_{ST}) calculated for the Portuguese Jews, Miranda do Douro and Northern Portuguese reference population based on a set of 15 autosomal markers, showed no statistically significant values after correction for multiple tests (data not

shown). The greatest distance was between the Bragança Jews and Miranda do Douro (F_{ST} = 0.00455, P = 0.02653); and the lowest between Miranda do Douro and the northern Portuguese population (F_{ST} = 0.00109, P = 0.23938). Pairwise F_{ST} based on the 9 autosomal shared STRs for all the Jewish and non-Jewish populations analyzed are represented in an MDS plot (Supplementary Figure S3).

X chromosome

Haplotype frequency data for the 12 X-chromosomal markers studied in the Portuguese Jews are shown in supplementary Table S2.

Pairwise LD analysis between pairs of loci (data not shown) revealed statistically significant allelic association between four pairs, namely DXS8378-DXS10135, DXS8378-DXS10148 and DXS10135-DXS10148 included in the linkage group 1 and DXS10101-DXS10103 contained within linkage group 3.

Values of average gene diversities across loci in the Bragança Jews (0.797±0.120) shows the lowest observed value, compared to the reference non Jewish populations, with a range that varies from a minimum detected in Miranda do Douro (0.807±0.100) (Pinto et al. 2014) to a maximum in the general Portuguese population (0.825±0.091) (Gomes et al. 2012). Considering the 4 linkage groups of the X chromosomal STR analyzed (Group 1 - DXS10148, DXS10135, DXS8378; Group 2 - DXS7132, DXS10079, DXS10074; Group 3 - DXS10103, HPRTB, DXS10101; and Group 4 - DXS10146, DXS10134, DXS7423), haplotype diversities found are summarized in Table 2.

In a locus-by-locus analysis (data not shown), the highest heterozygosity (0.928) was detected at locus DXS10135 and the lowest value (0.568) at locus DXS7423, as it was also found previously for the same loci in the population of Miranda do Douro and Zamora (Pinto et al. 2014). The general Portuguese population showed, however, its highest value at locus HPRTB (0.943) while the less diverse marker is DXS10135 (0.685).

All haplotypes observed in the Portuguese Jews as well as in the host population considered here (Miranda do Douro, Zamora and Portugal), were unique. Therefore, an indirect estimate locus by locus concerning the number of allele losses for the twelve

analyzed X-chromosomal markers was addressed. The results are shown in Supplementary Figure S4.

Pairwise F_{ST} genetic distances were calculated between the Bragança Jews, the Portuguese host populations and also the Spanish neighboring population of Zamora and the results are presented in Table 3. Significant genetic distances were obtained between the Portuguese Jews and all the other reference non-Jewish populations (Portugal, Miranda do Douro and Zamora). The highest genetic distance was observed between the Bragança Jews and Miranda do Douro (F_{ST} = 0.014, P < 5E-6) and the lowest between the Bragança Jews and Zamora (F_{ST} = 0.007, P = 0.003) with a distance of 0.011 (P < 5E-6) for the Jewish population and the general Portuguese population.

Discussion

Jewish communities worldwide have been intensely studied both from an anthropological perspective as well as from a genetic point of view, but the majority of researches have focused on Jews of Ashkenazim origin (Ostrer 2001) when compared with the few devoted to the small communities of Sephardic Jews. Uniparental genetic markers showed differences between and among these two groups concerning patterns of genetic diversity. Relative high levels of diversity were found in Sephardic Jews (Adams et al. 2008; Nogueiro et al. 2010; Nogueiro et al. 2014; Picornell et al. 2006; Shen et al. 2004), although some communities showed low diversities, as in (Behar et al. 2008), contrasting with rather strong founder effects and bottlenecks described in the Ashkenazi group (Behar et al. 2004; Behar et al. 2006; Behar et al. 2008; Ostrer and Skorecki 2013). While a significant sub-structure among Jewish populations was quite obvious from the study of these genetic markers, regarding recombinant markers, population differentiations were not so evident.

Concerning STR autosomal data, the available information on Jewish populations includes manly Sephardic Mediterranean populations (Picornell et al. 2003; Picornell et al. 2002). Among these communities and the Portuguese Jews, no significant differences in genetic diversities were found, either between the Jewish communities or between them and their respective non-Jewish host populations. Mean diversities were of the same range of magnitude (80%-82%) in most Jewish communities and each respective host populations with the exception of Iran and Libya that showed slightly

greater differences. Though the overall diversities are fairly high for communities historically known as closed and small, when the number of allele losses was evaluated, (Figure 1 and Figure S2), a reduction of diversity per locus was observed for the majority of markers studied for the Jewish communities when compared with the corresponding non-Jewish host populations. With the exception of very few locus, namely D7S820, D21S11, FGA, D3S1358, D13S317, D18S51 and D5S818 in Iran, Morocco, Turkey and the community of Mallorca, where the number of alleles was greater in the Jewish community, the tendency in all other communities was to maintain the same number of alleles as their host populations or much more frequently, to lose alleles. These results could be due to genetic drift, non-random mating strategies, and inbreeding and/or population size fluctuations and migration. Regarding the Portuguese Jews, with the exception of vWA locus, also a tendency to preserve or in the main, to lose alleles can be observed. Historically, while the Mediterranean Jewish communities could be to some extent considered rather open to their non-Jewish neighbors, at least till the 15th century, it is also known that after the decrees of expulsion and the Inquisition persecutions the Jewish reality changed abruptly (Barnavi 1992; Martins 2006). Many Iberian Jews fled to other countries, particularly in the Mediterranean, as Morocco or Turkey and as a consequence, particularly the Portuguese Jewish population was drastically reduced. Those who stayed became New-Christians, and some Crypto-Jews. This new imposed status associated with an increased social and religious pressure lead to a significant isolation of these Jewish communities (Barnavi 1992; Mea 1996; Mea 1997; Roth and Spivak 1946; Saraiva et al. 2001).

The genetic diversities found for the X chromosome are of the same magnitude as those found for autosomes in the Portuguese Jewish population. In comparison to the non-Jewish host populations, the haplotype diversities found for the X chromosome, are also very similar (Table 2) with slightly lower values in linkage groups 3 and 4. The only available data on other Jewish populations (Ferragut et al. 2015), nevertheless, shows much greater values than those observed in the present work. Since Ferragut et al. (2015) work pools together several distinct Jewish communities, such as Middle Eastern, Ashkenazi, Sephardic, North African and Chuetas, a direct comparison with those who share a more closely related history to the Portuguese Jews, namely the Chueta community from Mallorca Islands, is not possible.

Portuguese Jews showed nevertheless, quite high levels of genetic diversity for such a small and isolated community. A reduction in the number of alleles, less significant than that observed for the autosomes (Figure 1 and Supplementary Figure S2), was also detected for four (DXS10079, HPRTB, DXS10146 and DXS10134) out of the twelve X chromosome loci studied (Supplementary Figure S4).

Although LD is more likely to occur between physically close loci, it can also be an outcome of random genetic drift or admixture. Due to its haploidicity and absence of recombination in males, X chromosome can retain higher levels of LD for longer periods of time when compared to autosomes (Schaffner 2004). When association tests were carried out for the 12 X chromosome markers, significant allelic association was detected between four pairs of loci in the Portuguese Jews, two pairs in Zamora and one pair in Miranda do Douro and Portugal (Gomes et al. 2012; Pinto et al. 2014). Higher levels of LD were therefore detected in the Jewish population compared to the reference non-Jews. Nevertheless, in all four populations the allelic associations were always observed for pairs of loci within the same linkage group, as already reported elsewhere (Pinto et al. 2014; Poulsen et al. 2015; Zidkova et al. 2014).

Pairwise genetic distances (F_{ST}) were calculated for autosomal as well as for X-chromosomal markers. For the 9 common autosomal STRs studied including Jews and non-Jews, the MDS plot results (Supplementary Figure S3), showed no evident clustering of populations within any Mediterranean geographical context or a religious/ethnical background; only the Jewish community from Iran appearing as an outlier.

Concerning both autosomes and the X chromosome, since the resulting distances are based on the same kind and a similar number of markers, a comparison between the respective F_{ST} values can be done. In this comparison only the Portuguese Jews and the Portuguese host populations were considered since no data was available for such analysis in the other populations. In the results obtained for the 15 autosomal and the 12 X-chromosomal markers, the differentiation between Jews and non-Jews was approximately three times higher for the X chromosome comparatively to the autosomes.

Theoretically, if only random genetic drift was responsible for the differences between the populations of Jews and non-Jews, then it would be expected that the differentiation

observed in autosomes would equal 3/4 of that observed for the X chromosome. However, the value obtained for this genetic distance is much lower. The ratio between autosomes and the X chromosome is significantly lower than the above expectation. Although purging selection on males can be argued as a possible reason for such differences (Nam et al. 2015), the low value here reported suggests a sex biased effective population size over time in the Jewish population. Our previous works on uniparental markers for this community (Nogueiro et al. 2010; Nogueiro et al. 2014; Teixeira et al. 2011), showed the existence of Sephardic signatures as well as signs of introgression from the host non-Jewish population. Historically, and by Jewish religious law, Jewishness is an attribute transmitted throughout the maternal line. So, to some extent, this fact could influence a preference for a genetic flow from non-Jewish males into the genetic pool of the Portuguese Jews in detriment of a female mediated one. This demographic sex biased gene flow could explain the difference from the expected 3/4 ratio between the autosomes and the X chromosomes.

In conclusion, our results demonstrate the relevance of using different types of genetic markers with distinct inheritance patterns in the reconstruction and the analysis of the genetic portrait of populations with a complex history, such as the Portuguese Jews. The systematic comparison of differentiation patterns between autosomes and X chromosome in other Jewish groups, particularly Sephardic, and respective non-Jewish host populations may contribute in the future, for a more detailed reconstruction of the history of this heterogeneous and complex group. Especially interesting in this context of the intricate demographic history of the Sephardic diaspora would be the answer to the question if the gender biased gene flow reported here for the Portuguese Jews is a general trend of these communities. Genome-wide studies will surely contribute to the definition of much more detailed relations among different populations, and their demographic histories, namely if the powerful data sets they provide are analyzed in the framework here exemplified.

Acknowledgements

This work was supported by FCT, Portuguese Foundation for Science and Technology, financed by the European Social Funds (COMPETE-FEDER), and project PTDC/ATP-DEM/4545/2012. LA (SFRH/BPD/65000/2009), IN (SFRH/BD/73336/2010) are supported by FCT fellowships and funded by POPH-QREN – Promotion of scientific

employment, supported by the European Social Fund and national funds of the Ministry of Education and Science. IPATIMUP is an Associate Laboratory of the Portuguese Ministry of Science, Technology and Higher Education and is partially supported by FCT.

Literature cited

IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.

Adams SM, Bosch E, Balaresque PL, Ballereau SJ, Lee AC, Arroyo E, Lopez-Parra AM, Aler M, Grifo MS, Brion M et al. . 2008. The genetic legacy of religious diversity and intolerance: paternal lineages of Christians, Jews, and Muslims in the Iberian Peninsula. American journal of human genetics 83(6):725-736.

Amorim A, Alves C, Gusmão L, and Pereira L. 2006. Extended Northern Portuguese database on 21 autosomal STRs used in genetic identification. International Congress Series: Elsevier. p 364-366.

Arbiza L, Gottipati S, Siepel A, and Keinan A. 2014. Contrasting X-linked and autosomal diversity across 14 human populations. The American Journal of Human Genetics 94(6):827-844.

Balaresque PL, Ballereau SJ, and Jobling MA. 2007. Challenges in human genetic diversity: demographic history and adaptation. Human molecular genetics 16(R2):R134-R139.

Barnavi É. 1992. História Universal dos Judeus. Lisboa.: Contexto.

Barni F, Berti A, Pianese A, Boccellino A, Miller MP, Caperna A, and Lago G. 2007. Allele frequencies of 15 autosomal STR loci in the Iraq population with comparisons to other populations from the middle-eastern region. Forensic science international 167(1):87-92.

Behar DM, Hammer MF, Garrigan D, Villems R, Bonne-Tamir B, Richards M, Gurwitz D, Rosengarten D, Kaplan M, Della Pergola S et al. . 2004. MtDNA evidence for a genetic bottleneck in the early history of the Ashkenazi Jewish population. Eur J Hum Genet 12(5):355-364.

Behar DM, Metspalu E, Kivisild T, Achilli A, Hadid Y, Tzur S, Pereira L, Amorim A, Quintana-Murci L, Majamaa K et al. . 2006. The matrilineal ancestry of Ashkenazi Jewry: portrait of a recent founder event. American journal of human genetics 78(3):487-497.

Behar DM, Metspalu E, Kivisild T, Rosset S, Tzur S, Hadid Y, Yudkovsky G, Rosengarten D, Pereira L, Amorim A et al. . 2008. Counting the founders: the matrilineal genetic ancestry of the Jewish Diaspora. PLoS ONE 3(4):e2062.

Bonferroni CE. 1936. Teoria statistica delle classi e calcolo delle probabilita: Libreria internazionale Seeber.

Bulbul O, Fernandez-Formoso L, Phillips C, Altuncul H, Filoglu G, Lareu M, and Carracedo A. 2014. Allele frequencies of the five new European Standard Set (ESS) STRs and 15 established STRs in a Turkish population. Forensic Science International: Genetics 9:e26.

Coudray C, Guitard E, Keyser-Tracqui C, Melhaoui M, Cherkaoui M, Larrouy G, and Dugoujon J-M. 2007. Population genetic data of 15 tetrameric short tandem repeats (STRs) in Berbers from Morocco. Forensic science international 167(1):81-86.

Emery LS, Felsenstein J, and Akey JM. 2010. Estimators of the human effective sex ratio detect sex biases on different timescales. The American Journal of Human Genetics 87(6):848-856.

Excoffier L, and Lischer HE. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular ecology resources 10(3):564-567.

Ferragut J, Castro J, Ramon C, and Picornell A. 2015. Genetic diversity of 12 X-chromosomal short tandem repeats in Jewish populations. Forensic Science International: Genetics Supplement Series.

Garrigan D, and Hammer MF. 2006. Reconstructing human origins in the genomic era. Nature Reviews Genetics 7(9):669-680.

Gomes C, Magalhães M, Alves C, Amorim A, Pinto N, and Gusmão L. 2012. Comparative evaluation of alternative batteries of genetic markers to complement autosomal STRs in kinship investigations: autosomal indels vs. X-chromosome STRs. International journal of legal medicine 126(6):917-921.

Hammer MF, Mendez FL, Cox MP, Woerner AE, and Wall JD. 2008. Sex-biased evolutionary forces shape genomic patterns of human diversity. PLoS genetics 4(9):e1000202.

Heyer E, Chaix R, Pavard S, and Austerlitz F. 2012. Sex-specific demographic behaviours that shape human genomic variation. Molecular ecology 21(3):597-612.

Jobling MA. 2012. The impact of recent events on human genetic diversity. Philosophical transactions of the Royal Society of London 367(1590):793-799.

Jobling MA, and Tyler-Smith C. 2003. The human Y chromosome: an evolutionary marker comes of age. Nature reviews 4(8):598-612.

Karafet TM, Mendez FL, Meilerman MB, Underhill PA, Zegura SL, and Hammer MF. 2008. New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. Genome research 18(5):830-838.

Khodjet-el-Khil H, Fadhlaoui-Zid K, Gusmao L, Alves C, Benammar-Elgaaied A, and Amorim A. 2012. Allele frequencies for 15 autosomal STR markers in the Libyan population. Annals of human biology 39(1):80-83.

Khodjet-El-Khil H, Fadhlaoui-Zid K, Gusmão L, Alves C, Benammar-Elgaaied A, and Amorim A. 2008. Substructure of a Tunisian Berber population as inferred from 15 autosomal short tandem repeat loci. Human biology 80(4):435-448.

Martins J. 2006. Portugal e os judeus: Judaísmo e anti-semitismo no século XX. Lisboa: Vega.

Mea E. 1996. O Judaísmo no Séc. XX - A fénix renascida em Trás-os-Montes. Congresso Histórico 450 Anos da Fundação da Diocese de Bragança-Miranda. Bragança: Fundação da Diocese de Bragança-Miranda.

Mea EA. 1997. A inquisição de Coimbra no século XVI: a instituição, os homens e a sociedade: Fundação Engo António de Almeida.

Nam K, Munch K, Hobolth A, Dutheil JY, Veeramah KR, Woerner AE, Hammer MF, Mailund T, Schierup MH, and Prado-Martinez J. 2015. Extreme selective sweeps independently targeted the X chromosomes of the great apes. Proceedings of the National Academy of Sciences:201419306.

Nogueiro I, Manco L, Gomes V, Amorim A, and Gusmao L. 2010. Phylogeographic analysis of paternal lineages in NE Portuguese Jewish communities. American journal of physical anthropology 141(3):373-381.

Nogueiro I, Teixeira J, Amorim A, Gusmao L, and Alvarez L. 2014. Echoes from Sepharad: signatures on the maternal gene pool of crypto-Jewish descendants. Eur J Hum Genet.

Nogueiro IP, Teixeira J, Amorim A, Gusmão L, and Alvarez L. 2015. PORTUGUESE CRYPTO-JEWS: THE GENETIC HERITAGE OF A COMPLEX HISTORY. Frontiers in Genetics 6:12.

Ostrer H. 2001. A genetic profile of contemporary Jewish populations. Nature reviews 2(11):891-898.

Ostrer H, and Skorecki K. 2013. The population genetics of the Jewish people. Human genetics 132(2):119-127.

Picornell A, Gimenez P, Castro J, and Ramon M. 2006. Mitochondrial DNA sequence variation in Jewish populations. International journal of legal medicine 120(5):271-281.

Picornell A, Tomàs C, Castro J, and Ramon M. 2003. STR data for 13 loci from Jewish populations. International Congress Series: Elsevier. p 109-112.

Picornell A, Tomas C, Jimenez G, Castro JA, and Ramon MM. 2002. Jewish population genetic data in 20 polymorphic loci. Forensic science international 125(1):52-58.

Pinto J, Pereira V, Marques S, Amorim A, Alvarez L, and Prata M. 2014. Mirandese language and genetic differentiation in Iberia: a study using X chromosome markers. Annals of human biology 42(1):20-25.

Poulsen L, Farzad MS, Børsting C, Tomas C, Pereira V, and Morling N. 2015. Population and forensic data for three sets of forensic genetic markers in four ethnic groups from Iran: Persians, Lurs, Kurds and Azeris. Forensic Science International: Genetics 17:43-46.

Roth C, and Spivak A. 1946. Historia de los marranos: Editorial Israel.

Saraiva AJ, Salomon HP, and Sassoon ISD. 2001. The Marrano Factory: The Portuguese Inquisition and Its New Christians 1536-1765. Leiden: Brill.

Schaffner SF. 2004. The X chromosome in population genetics. Nature Reviews Genetics 5(1):43-51.

Shen P, Lavi T, Kivisild T, Chou V, Sengun D, Gefel D, Shpirer I, Woolf E, Hillel J, Feldman MW et al. . 2004. Reconstruction of patrilineages and matrilineages of Samaritans and other Israeli populations from Y-chromosome and mitochondrial DNA sequence variation. Human mutation 24(3):248-260.

Shepard E, and Herrera R. 2006. Iranian STR variation at the fringes of biogeographical demarcation. Forensic science international 158(2):140-148.

Takezaki N, Nei M, and Tamura K. 2010. POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. Molecular biology and evolution 27(4):747-752.

Team RC. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2012. ISBN 3-900051-07-0.

Teixeira JC, Nogueiro I, Goios A, Gusmão L, Amorim A, and Alvarez L. 2011. Mitochondrial DNA-control region sequence variation in the NE Portuguese Jewish community. FSI genetics supplement series 3(1):e51-e52.

Tomàs C, Picornell A, Castro J, Ramon M, Gusmao L, Lareu M, and Carracedo A. 2000. Genetic variability at nine STR loci in the Chueta (Majorcan Jews) and the Balearic populations investigated by a single multiplex reaction. International journal of legal medicine 113(5):263-267.

Underhill PA, and Kivisild T. 2007. Use of y chromosome and mitochondrial DNA population structure in tracing human migrations. Annu Rev Genet 41:539-564.

van Oven M, and Kayser M. 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Human mutation 30(2):E386-394.

Zegura SL. 2005. Human Evolutionary Genetics: Origins, Peoples and Disease. American journal of human genetics 76(6):1087.

Zidkova A, Capek P, Horinek A, and Coufalova P. 2014. Investigator® Argus X-12 study on the population of Czech Republic: Comparison of linked and unlinked X-STRs for kinship analysis. Electrophoresis 35(14):1989-1992.

Unveiling the Sephardic Portuguese genetic legacy from autosomal and X chromosome STRs markers – Figures, Tables and supplementary material

Figure S1. Geographical location of the Portuguese and Spanish populations considered in the present work.

TABLE S1. Allele frequencies for 15 autosomal STR loci in a sample of 53 unrelated Portuguese Jews

Allele	D2S1338	D3S1358	D5S818	D7S820	D8S1179	D13S317	D16S539	D18S51	D19S433	D21S11	FGA	TH01	TPOX	vWA	CSF1PO
5	-	-	-	-	-	-	-	-	-	-	-	0,009	-	-	-
6	-	-	-	-	-	-	-	-	-	-	-	0,245	-	-	-
7	-	-	-	-	-	-	-	-	-	-	-	0,057	-	-	-
8	-	-	-	0,142	0,019	0,094	0,009	-	-	-	-	0,160	0,462	-	-
9	-	-	0,047	0,132	-	0,104	0,094	-	-	-	-	0,226	0,132	-	0,047
9,3	-	-	-	-	-	-	-	-	-	-	-	0,283	-	-	-
10	-	-	0,047	0,274	0,104	0,038	0,057	-	-	-	-	0,019	0,113	-	0,264
11	-	-	0,321	0,274	0,113	0,349	0,443	0,009	-	-	-	-	0,264	-	0,226
12	-	-	0,425	0,142	0,104	0,255	0,255	0,113	0,118	-	-	-	0,028	-	0,377
13	-	-	0,151	0,038	0,330	0,085	0,123	0,142	0,245	-	-	-	-	-	0,075
13.2	-	-	-	-	-	-	-	-	0,010	-	-	-	-	-	-
14	-	0,075	0,009	-	0,057	0,075	0,019	0,170	0,196	-	-	-	-	0,066	0,009
14.2	-	-	-	-	-	-	-	-	0,020	-	-	-	-	-	-
15	-	0,274	-	-	0,236	-	-	0,151	0,176	-	-	-	-	0,142	-
15.2	-	-	-	-	-	-	-	-	0,118	-	-	-	-	-	-
16	0,066	0,292	-	-	0,038	-	-	0,160	0,049	-	-	-	-	0,245	-
16.2	-	-	-	-	-	-	-	-	0,069	-	-	-	-	-	-
17	0,264	0,226	-	-	-	-	-	0,123	-	-	-	-	-	0,226	-
18	0,094	0,085	-	-	-	-	-	0,066	-	-	-	-	-	0,217	-
19	0,132	0,047	-	-	-	-	-	0,019	-	-	0,066	-	-	0,047	-
20	0,151	-	-	-	-	-	-	0,038	-	-	0,179	-	-	0,057	-
21	0,028	-	-	-	-	-	-	-	-	-	0,189	-	-	-	-
22	0,028	-	-	-	-	-	-	-	-	-	0,245	-	-	-	-
23	0,085	-	-	-	-	-	-	-	-	-	0,057	-	-	-	-
24	0,132	-	-	-	-	-	-	0,009	-	-	0,123	-	-	-	-
25	0,009	-	-	-	-	-	-	-	-	-	0,075	-	-	-	-
26	0,009	-	-	-	-	-	-	-	-	-	0,019	-	-	-	-
27	-	-	-	-	-	-	-	-	-	-	0,047	-	-	-	-
28	-	-	-	-	-	-	-	-	-	0,198	-	-	-	-	-
29	-	-	-	-	-	-	-	-	-	0,189	-	-	-	-	-
30	-	-	-	-	-	-	-	-	-	0,217	-	-	-	-	-
30.2	-	-	-	-	-	-	-	-	-	0,057	-	-	-	-	-
31	-	-	-	-	-	-	-	-	-	0,047	-	-	-	-	-
31.2	-	-	-	-	-	-	-	-	-	0,142	-	-	-	-	-
32	-	-	-	-	-	-	-	-	-	0,009	-	-	-	-	-
32.2	-	-	-	-	-	-	-	-	-	0,085	-	-	-	-	-
33.2	-	-	-	-	-	-	-	-	-	0,057	-	-	-	-	-

TABLE 1. Genetic diversity standard indices for 15 STR loci in Bragança Jews, Miranda do Douro and Northern Portugal non-Jews

	Allele	D2S1338	D3S1358	D5S818	D7S820	D8S1179	D13S317	D16S539	D18S51	D19S433	D21S11	FGA	TH01	TPOX	vWA	CSF1PO
Bragança Jews (N=53)																
H_{OBS}		0,6981	0,6981	0,6981	0,8113	0,7547	0,6604	0,6415	0,8679	0,6078	0,8113	0,9057	0,7170	0,6226	0,7736	0,5660
H_{EXP}		0,8584	0,7806	0,6961	0,7989	0,8036	0,7867	0,7177	0,8767	0,8433	0,8501	0,8492	0,7865	0,6920	0,8194	0,7355
P_{HW}		0,0978	0,2342	0,2836	0,8952	0,2802	0,1140	0,1531	0,6337	0,0032	0,5150	0,7074	0,6972	0,0540	0,2141	0,0142
Miranda do Douro (N=87)																
H_{OBS}		0,7586	0,8161	0,7471	0,7356	0,8391	0,8023	0,7471	0,8506	0,7294	0,8140	0,8372	0,7241	0,6047	0,7586	0,7093
H_{EXP}		0,8629	0,7850	0,7010	0,7932	0,7889	0,7659	0,8149	0,8815	0,8308	0,8177	0,8627	0,8000	0,6317	0,8132	0,6960
P_{HW}		0,3299	0,4417	0,4768	0,2146	0,0967	0,8032	0,1061	0,2338	0,0054	0,0471	0,4504	0,3532	0,4973	0,1426	0,2064
Northern Portuguese (N=105)																
H_{OBS}		0,8000	0,8286	0,6571	0,8095	0,8191	0,7905	0,7524	0,8952	0,7429	0,8000	0,8667	0,7524	0,6857	0,7714	0,6857
H_{EXP}		0,8465	0,7879	0,6900	0,7954	0,8086	0,8004	0,7728	0,8863	0,7959	0,8600	0,8612	0,8041	0,6715	0,8026	0,7484
P_{HW}		0,1498	0,2318	0,4150	0,7747	0,2600	0,5710	0,9441	0,4570	0,3675	0,0711	0,5979	0,8710	0,2735	0,9561	0,1562

HOBS, observed heterozygosity; HEXP, expected heterozygosity; P HW, Hardy-Weinberg equilibrium P values.

Figure 1. Number of allelic losses for the studied populations considering nine loci. The black color represents the non-Jewish population of northern Portugal; dark-grey represents the population of Miranda do Douro and light-grey the Jewish population of Bragança.

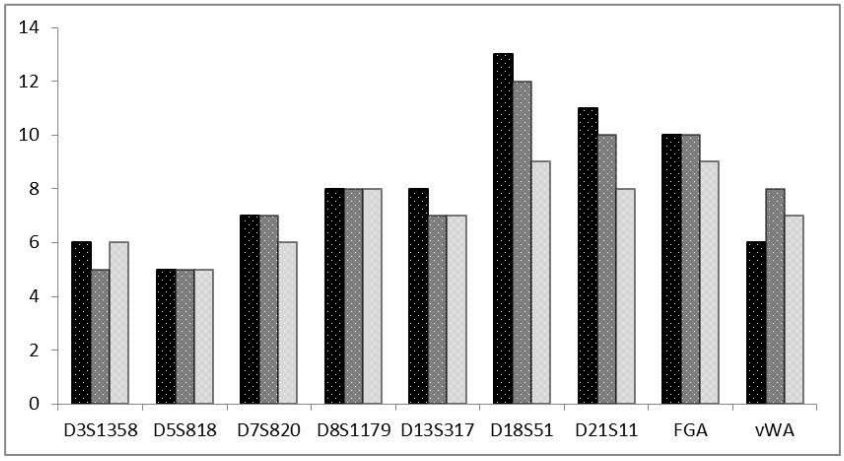
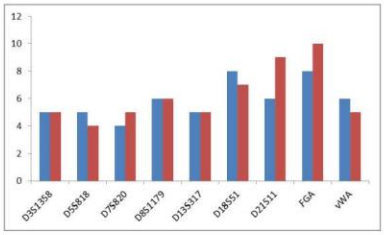
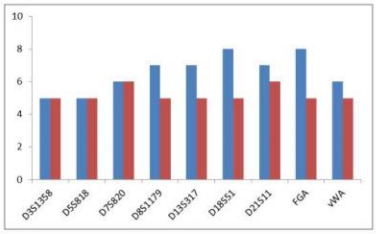


Figure S2. Number of allelic losses considering nine common loci. Red represents the Jewish population and Blue denotes the respective non-Jewish host populations.

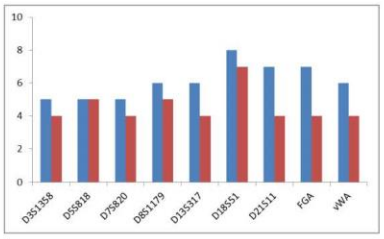
A) Iran



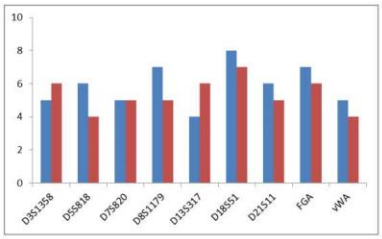
B) Iraq



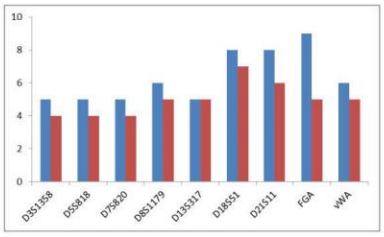
C) Libya



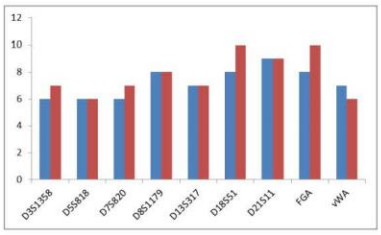
D) Morocco



E) Tunisia



F) Turkey



G) Mallorca Island

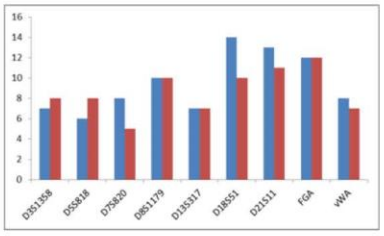


Figure S3. Multidimensional scaling analysis based on pairwise F_{ST} genetic distances calculated for 9 autosomal STRs (D3S1358, D5S818, D7S820, D8S1179, D13S317, D18S51, D21S11, FGA and vWA) for the Jewish populations and respective non-Jewish host populations (NJ).

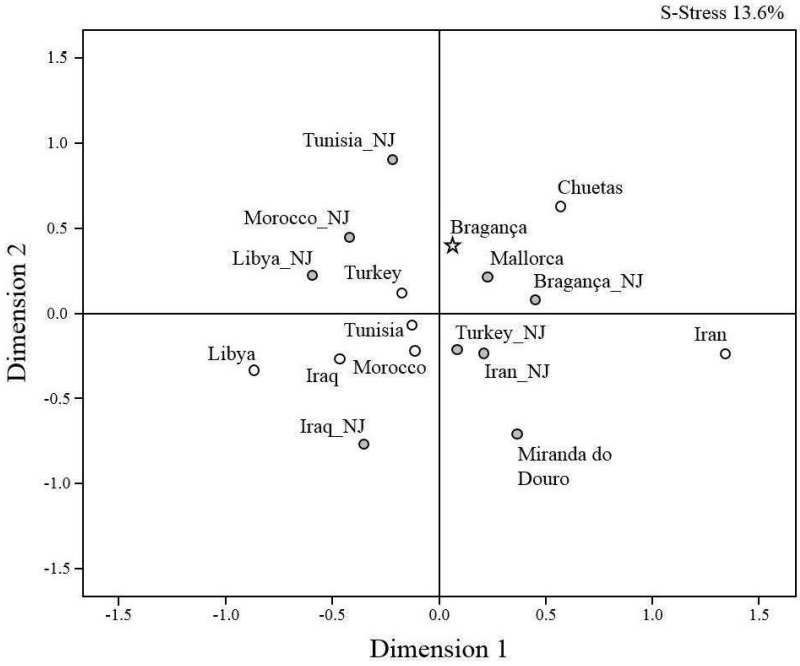


Table S2. Haplotype frequency data for the 12 X-chromosomal markers studied in the Portuguese Jews

X - STR	Haplotype	Frequency	s.d.
DXS10148	22.1 17 12	0.1481	0.0488
DXS10135	18 28 10	0.0185	0.0185
DXS8378	24.1 27 11	0.037	0.0259
	27.2 21 11	0.037	0.0259
	19 29 11	0.0185	0.0185
	28.1 23 11	0.0556	0.0315
	29.1 23 11	0.0185	0.0185
	24.1 26 12	0.0185	0.0185
	25.1 28 10	0.037	0.0259
	20.1 22 11	0.0185	0.0185
	26.1 27 10	0.0185	0.0185
	27.1 27 10	0.037	0.0259
	23.1 26 12	0.0185	0.0185
	23 29 11	0.0185	0.0185
	25.1 21.1 10	0.0185	0.0185
	23 26 11	0.0185	0.0185
	24.1 20.1 10	0.0185	0.0185
	27.1 27 11	0.0185	0.0185
	27.1 18 10	0.0185	0.0185
	19 26 11	0.0185	0.0185
	27.1 21 11	0.0185	0.0185
	24.1 23 12	0.037	0.0259
	24.1 18 10	0.0185	0.0185
	30.1 27 10	0.0185	0.0185
	23.1 24 10	0.0185	0.0185
	23 19 10	0.0185	0.0185
	23.1 24 11	0.0185	0.0185
	28.1 31 11	0.0185	0.0185
	26.1 19 12	0.0185	0.0185
	24 30 11	0.0185	0.0185
	23.1 25 11	0.0185	0.0185
	25.1 30 10	0.0185	0.0185
	24.1 22 12	0.0185	0.0185
	27.1 17 11	0.0185	0.0185
	18 30 11	0.0185	0.0185
	25.1 19 11	0.0185	0.0185
	24.1 23 10	0.0185	0.0185
	25.1 20 10	0.0185	0.0185
	18 18.1 11	0.0185	0.0185
	23 21 11	0.0185	0.0185

s.d. - standard deviation

Table S2. Haplotype frequency data for the 12 X-chromosomal markers studied in the Portuguese Jews

X - STR	Haplotype	Frequency	s.d.
DXS7132	12 18 18	0.018519	0.018519
DXS10079	12 18 8	0.055556	0.031464
DXS10074	16 18 16	0.037037	0.025941
	14 15 20	0.018519	0.018519
	12 20 8	0.037037	0.025941
	15 20 17	0.018519	0.018519
	14 20 16	0.055556	0.031464
	13 18 15	0.037037	0.025941
	14 19 17	0.018519	0.018519
	13 20 17	0.055556	0.031464
	13 17 7	0.018519	0.018519
	13 19 17	0.018519	0.018519
	14 16 17	0.037037	0.025941
	15 19 16	0.018519	0.018519
	15 21 16.2	0.018519	0.018519
	13 21 16	0.055556	0.031464
	15 19 8	0.018519	0.018519
	14 18 18	0.018519	0.018519
	15 16 18	0.018519	0.018519
	13 20 18	0.055556	0.031464
	15 21 17	0.037037	0.025941
	14 20 17	0.018519	0.018519
	16 19 15	0.018519	0.018519
	14 19 8	0.037037	0.025941
	15 18 8	0.018519	0.018519
	14 19 16	0.018519	0.018519
	13 16 17	0.018519	0.018519
	13 20 7	0.018519	0.018519
	12 19 19	0.018519	0.018519
	13 19 8	0.037037	0.025941
	12 17 17	0.018519	0.018519
	16 15 19	0.018519	0.018519
	15 18 15	0.018519	0.018519
	13 17 8	0.018519	0.018519
	14 19 9	0.037037	0.025941
	13 15 17	0.018519	0.018519
s.d. - standard deviation			

Table S2. Haplotype frequency data for the 12 X-chromosomal markers studied in the Portuguese Jews

X - STR	Haplotype	Frequency	s.d.
DXS10103	19 12 29.2	0.1667	0.0512
HPRTB	17 14 31	0.0556	0.0315
DXS10101	17 13 28.2	0.0185	0.0185
	19 11 28.2	0.037	0.0259
	19 12 31.2	0.0185	0.0185
	17 13 31	0.0185	0.0185
	19 12 28	0.037	0.0259
	19 12 30.2	0.0556	0.0315
	16 11 29.2	0.0185	0.0185
	18 12 33	0.0185	0.0185
	18 13 30.2	0.0185	0.0185
	17 12 30	0.0185	0.0185
	19 13 33.2	0.0185	0.0185
	20 13 34	0.0185	0.0185
	18 12 30.2	0.0185	0.0185
	18 14 32.2	0.0741	0.036
	20 12 30.2	0.037	0.0259
	18 12 30	0.037	0.0259
	16 12 32	0.0185	0.0185
	20 13 31.2	0.0185	0.0185
	18 13 30	0.0185	0.0185
	21 11 29.2	0.0185	0.0185
	16 12 31	0.037	0.0259
	20 13 32.2	0.0185	0.0185
	19 13 32.2	0.0185	0.0185
	17 13 28	0.0185	0.0185
	19 14 30.2	0.037	0.0259
	16 12 28.2	0.037	0.0259
	18 12 31.2	0.0185	0.0185
	20 12 31.2	0.0185	0.0185
	18 14 35	0.0185	0.0185
	19 14 32.2	0.0185	0.0185

s.d. - standard deviation

Table S2. Haplotype frequency data for the 12 X-chromosomal markers studied in the Portuguese Jews

X - STR	Haplotype	Frequency	s.d.
DXS10146	31 35 15	0.0556	0.0315
DXS10134	28 37 15	0.0556	0.0315
DXS7423	26 35 15	0.0556	0.0315
	28 36 14	0.0185	0.0185
	25 36 14	0.0185	0.0185
	43.2 37 15	0.0185	0.0185
	29 35 15	0.0741	0.036
	28 33 15	0.0185	0.0185
	38.2 35.3 15	0.0185	0.0185
	29 36 15	0.0185	0.0185
	27 37 15	0.0556	0.0315
	29 34 15	0.0185	0.0185
	38.2 40.3 17	0.037	0.0259
	39.2 40.3 15	0.037	0.0259
	43.2 34 16	0.0185	0.0185
	45.2 35 15	0.0185	0.0185
	30 38 15	0.0185	0.0185
	30 33 17	0.0185	0.0185
	25 36 13	0.0185	0.0185
	25 37 14	0.0185	0.0185
	31 36 15	0.037	0.0259
	39.2 38 15	0.0185	0.0185
	31 34 14	0.037	0.0259
	29 36 16	0.0185	0.0185
	31 35 16	0.0185	0.0185
	30 35 14	0.0185	0.0185
	39.2 39.3 15	0.0185	0.0185
	46.2 36 16	0.0185	0.0185
	45.2 36 15	0.0185	0.0185
	43.2 38.3 14	0.0185	0.0185
	38.2 42.3 17	0.0185	0.0185
	31 32 15	0.0185	0.0185
	45.2 33 16	0.0185	0.0185
	31 37 15	0.0185	0.0185
	27 35 15	0.037	0.0259
	43.2 38 14	0.0185	0.0185
	40.2 40.3 16	0.0185	0.0185
	28 35 14	0.0185	0.0185
s.d. - standard deviation			

Table 2. Haplotype diversities of the 4 X - chromosomal STR linkage groups for the Portuguese Jews and other Jewish and non-Jewish populations.

Population	Linkage group 1	Linkage group 2	Linkage group 3	Linkage group 4
Portuguese Jews (N=54)	0.835	0.812	0.764	0.778
Other Jews [a] (N=152)	0.997	0.992	0.988	0.995
Non-Jews				
Miranda [b] (N=55)	0.830	0.813	0.774	0.800
Zamora [c] (N=116)	0.841	0.794	0.788	0.806
Portugal [d] (N=136)	0.848	0.805	0.820	0.829

Linkage groups comprise the following loci, respectively:1: DXS10148-DXS10135-DXS8378; 2: DXS7132-DXS10079-DXS10074; 3: DXS10103-HPRTB-DXS10101; 4: DXS10146-DXS10134-DXS7423; a) Ferragut et al. (2015); b) Pinto et al. (2014); c) Pinto et al. (2014); Gomes et al. (2012).

Table 3. Pairwise F_{ST} genetic distances (below the diagonal) and respective p values (above the diagonal) obtained between the Bragança Jewish population and the geographical non Jewish host populations from Zamora, Miranda do Douro and Portugal for 12 X-STRs.

FST	Zamora	Miranda do Douro	Portugal	Bragança Jews
Zamora	*	0.000	0.005	0.003
Miranda do Douro	0.005	*	0.002	0.000
Portugal	0.002	0.003	*	0.000
Bragança Jews	0.007	0.014	0.011	*

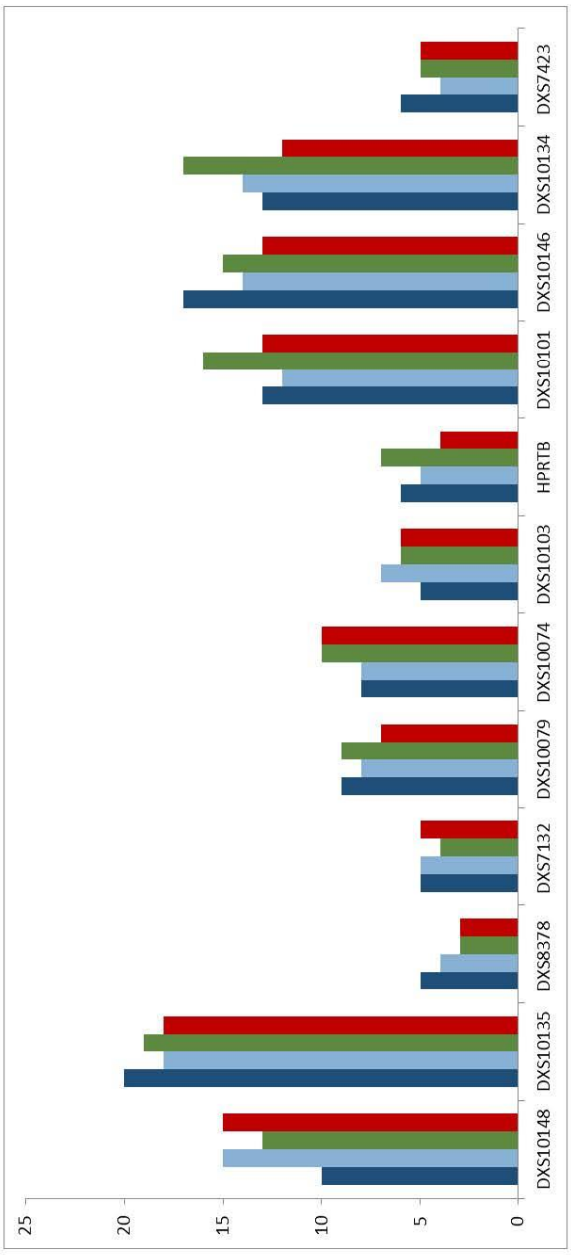


Figure S4. Number of allelic losses for the studied populations considering the twelve X chromosome markers. The dark-blue color represents the population of Zamora; light-blue for Miranda do Douro; green for Portugal and red represents the Jewish population of Bragança.

CHAPTER IV: PRELIMINARY RESULTS OF THE DIASPORA JEWS IN BRAZIL

Uniparental lineages of the Jewish Diaspora populations` in Brazil

Introduction

The history of the Jews in Brazil can be considered a unique case, since the Jews have lived there throughout the very beginning of the country, contributing substantially to its economic and social development [151]. As it was previously mention under the topic "**Jewish Diaspora in Brazil"** of this thesis, Jewish history in Brazil comprehends several distinct phases, with periods of great expansion and prosperity alternating with periods of anti-Judaic discrimination. As a result, similarly to what happened in Portugal, the phenomenon of Crypto-Judaism resurfaced in Brazil in the 20th-21st centuries, particularly in the Northeast Brazil, where some are claiming an unconditional return to Judaism due to the awareness of their Jewish ancestry [24, 25]. It is important to note that the Brazilian New-Christians or Crypto-Jews have nevertheless, original features that distinguish them clearly from the New-Christians who migrated to the more uniform and segregating societies of Europe: the extensive admixture with the native population, creating deep roots in the new land and a full integration among the social and political organization as well as the development of a diversity of attitudes and behaviors tremendously different from their original background [23, 24].

As an extremely heterogeneous population that gathers Amerindian, African and European ancestries, the Brazilians have been deeply scrutinized for their maternal [152-158] and paternal [159-165] genetic lineages. In these studies, both African and European contributions for the present Brazilian genetic pool are addressed. Nevertheless, with the exception of very few research papers, two of them concerning the phylogeography of Brazilian Y-chromosome lineages [162, 163] and a recent one focused on Laron syndrome (a genetic disorder caused by mutations in the growth hormone receptor gene [166]), no systematic study has yet been accomplished on the Brazilian Diaspora Jewish population. Thus the present study constitutes a first draft of the uniparental lineages of the Brazilian Jews, particularly of those with Iberian ancestry. In the following, a total of 37 individuals from the Jewish communities of Rio de Janeiro, Recife and Natal were analyzed (**Figure 4**). The fundamental aims of this work are the clarification as to what extent they kept not only their cultural identity, but also their genetic make-up; if the cultural isolation has implied the impoverishment of their genetic diversity and if their genetic profile reveals the incorporation of host population contributions. In this study, we

begin reporting the first answers to these questions for the male and female gene pool using Y-chromosome SNP markers and sequencing the complete control region of the mtDNA.

Figure 4. Geographic location of the sampling area in the Brazilian context.

Material and Methods

Sample collection

A total of 37 samples were collected in the Jewish community of Rio de Janeiro (N=9), Recife (N=7) and Natal (N=21). Religious, ethno-historical, and individual affiliation criteria were used to select DNA samples. Personal inquiries were made to each individual, to confirm their Jewish ancestry and all samples were taken under informed consent. The study was approved by the Ethics Committee of Hospital Universitário Pedro Ernesto/UERJ (N 923.816).

DNA analysis

Y Chromosome

Twenty Y chromosome binary markers were determined (12f2a, 92R7, M9, M13, M17, M18, M26, Tat, M62, M70, M78, M81, M123, M170, M172, M173, M201, M213, P25 and SRY1532) by SNaPshot technique using 4 multiplex assays [167, 168]. Furthermore, the M269 Y-SNP was typed by RFLP technique [169]. For those Y chromosomes carrying the M213 ancestral status

(therefore excluded from macro-haplogroup F) the DE haplogroup clade was tested through the analysis of M1 (YAP) indel polymorphism, screened by a single PCR reaction with direct visualization in a polyacrylamide gel [170]. The samples with a M269 derived status were also analyzed for 7 additional downstream Y-SNPs (L23, M153, M167, M529, S116, U106 and U152) (Resque et al., personal communication), using a SNaPshot method according to the manufacturers' recommendations (Applied Biosystems).

mtDNA

A 3348 bp fragment containing the entire mtDNA-CR was amplified in two overlapping fragments (14898-151/16488-1677) by polymerase chain reaction (PCR), using mtDNA specific primers [171].

Data Analysis

Haplogroup frequencies were calculated through direct counting. Standard diversity indices were estimated using ARLEQUIN v3.5.1.2 software [172]. Y-Chromosomal Haplogroup nomenclature is according to van Oven et al. (2014) [112]. Mitochondrial haplogroups were assigned using Haplogrep [173] and EMMA [174] software, following the updated mtDNA phylogeny, PhyloTree, mtDNA tree Build 16 [138].

Results and discussion

Y Chromosome

The 29 SNPs typed in this work allowed for the discrimination of 10 different haplogroups represented in **Figure 5**, along with the corresponding frequencies for each of the three Jewish communities. In the total sample of the Brazilian Jews, haplogroups R1b1a – S116*, J2 – M172 and G – M201 are the most frequent lineages, adding up to 62.5% (**Figure 6**). The remaining seven haplogroups occur with much lower frequency, in most cases representing only one individual. Overall, Y chromosome haplogroup composition is quite similar to that previously reported for Brazilian populations [162, 163, 175].

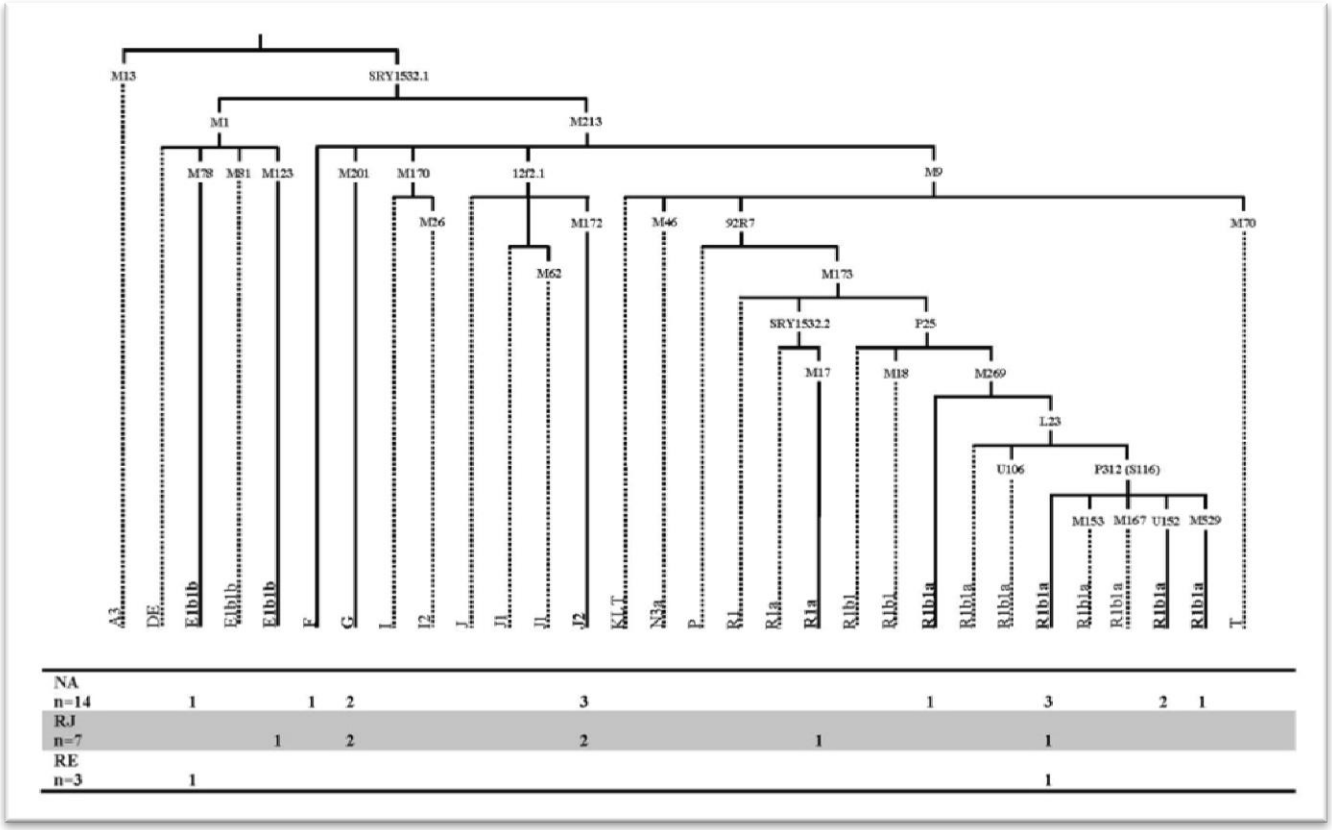


Figure 5. Phylogenetic tree of all the haplogroups tested and corresponding absolute frequencies for the communities of Natal (NA), Rio de Janeiro (RJ) and Recife (RE). Haplogroups detected in this work are highlighted in bold.

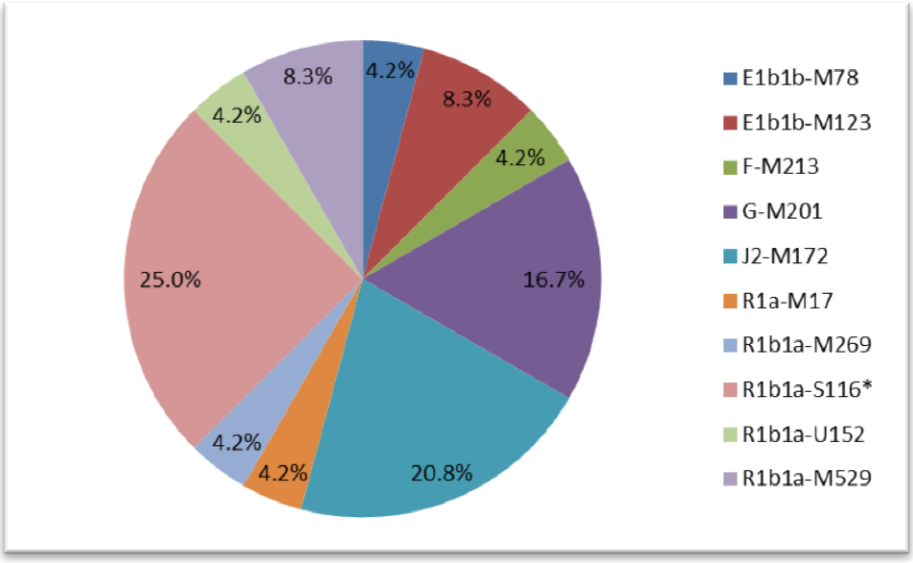


Figure 6. Y chromosome haplogroup composition and corresponding relative frequencies for the total sample of the Brazilian Jewish communities.

Regarding R-M269*, the Iberian P312 (S116), is the most frequent (25%) sub-lineage found among the Brazilian Jewish sample, in accordance to what was previously found for the Portuguese population, including the Portuguese Jews (Marques et al, personal communication), and also for the general Brazilian populations (Resque et al., personal communication). Haplogroups J2 – M172 and G – M201, also appear with similar frequencies among other populations in Brazil [175]. These two lineages were also detected in Iberian Jews with high frequencies [176].

The autochthonous Native American Q1a3a* (with all its sub lineages Q1a3a1, Q1a3a2, Q1a3a3 and Q1a3a4) was absent in the Jewish sample as well as lineages with sub-Saharan origin.

Analyses at the intra-population level revealed a relatively high value for the expected heterozygosity (0.8804 ± 0.0387) compared to the value previously reported for the general population of Rio de Janeiro (0.7195 ± 0.0413) [165]. However, in this study, the sub-lineages of the R-M269 clade were not discriminated. In fact, in a larger sample including all five geopolitical regions of Brazil, where the same level of resolution such as that of the present work was used, a similar, although slightly lower, mean value of expected heterozygosity (0.8463±0.00521) was found (Resque et al., personal communication). Indeed, when haplotype diversities were re-calculated pooling R1b1a-M269 sub-lineages, a significantly lower value of haplogroup diversity was observed in the Jewish sample (0.7754 ± 0.0633).

Mitochondrial DNA

With the analysis of 37 complete Control Region (CR) sequences it was possible to identify 17 different mitochondrial haplotypes. Detailed information on haplotype composition is available on **Table 2**. These 37 samples allowed the definition of 10 different main haplogroups. Respective relative frequencies are shown in **Figure 7**.

In general, haplogroup frequencies found in the Jewish sample, are in accordance with what is expected for a typical admixed Brazilian population, namely due to the frequency values obtained for the major Amerindian haplogroups (A= 11%; B=19%; C=14% and D=3%), sub-Saharan Africa (L=19%) and European (H=16%; HV0=5%; U5=3% and V=3%).

Table 2. Detailed information of haplogroup and haplotype composition of the Brazilian Jewish community: Rio de Janeiro (RJ), Natal (Na) and Recife (RE).

Sample ID	CR Mutated Positions	Haplogroup
RJ1	16223T,16325C,16327T,87.1G,89C,146C,249del,263G,290del,291del,309.1C,315.1C,489C,493G,523del,524del	C1b
RJ2	16153A,16298C,72C,93G,263G,309.1C,309.2C,315.1C,316A	V7a
RJ3	16248T,16519C,263G,309.1C,309.2C,315.1C	H3w
RJ4	16129A,16354T,263G,309.1C,315.1C	H2a1
RJ5	16189C,16217C,16519C,73G,263G,315.1C,499A,524.1A,524.2C	B4b
RJ6	16189C,16223T,16278T,16294T,16300G,16309G,16390A,16519C,73G,143A,146C,152C,195C,263G,264T,315.1C,510T,534T	L2a1
RJ7	16188.1C,16223T,16248A,16298C,16325C,16327T,16360T,73G,152C,249del,263G,290del,291del,309.1C,315.1C,489C,493G,523del,524del	C1b
RJ8	16126C,16187T,16189C,16223T,16264T,16270T,16278T,16311C,16519C,73G,152C,182T,185T,195C,247A,263G,315.1C,357G,523del,524del	L1b
RJ9	16051G,16182C,16183C,16189C,16217C,16519C,73G,152C,263G,309.1C,309.2C,315.1C,499A	B2b+152
Na01	16183C,16189C,16217C,16241G,16519C,73G,103A,152C,263G,315.1C,499A	B2+152
Na02	16189C,16519C,263G,315.1C	H1+16189
Na03	16092C,16111T,16223T,16290T,16319A,16362C,16519C,64T,73G,146C,152C,235G,263G,295G,309.1C,315.1C,523del,524del	A2+(64)+@153
Na04	16147T,16223T,16298C,16325C,16327T,16400T,73G,249del,263G,290del,291del,315.1C,489C,493G,523del,524del	C1b
Na05	16111T,16182M,16183C,16189C,16223T,16290T,16319A,16362C,64T,73G,146C,153G,200G,215G,235G,263G,309.1C,309.2C,315.1C,523del,524del	A2+(64)+16189
Na06	16104T,16223T,73G,152C,153G,182T,200G,263G,315.1C,523del,524del	X1
Na07	16187A,16223T,16325C,16362C,73G,263G,315.1C,489C	D1
Na10	16178C,16183C,16189C,16217C,16519C,73G,263G,309.1C,309.2C,315.1C,499A	B4b
Na11	16104T,16111T,16223T,16290T,16319A,16362C,16519C,64T,73G,146C,153G,235G,263G,309.1C,309.2C,315.1C,523del,524del,573.1N	A2+(64)
Na12	16129A,16183C,16189C,16193.1C,16215G,16223T,16278T,16294T,16311C,16355T,16360T,16519C,73G,151T,152C,182T,183G,186A,189C,204C,247A,257G,263G,315.1C,316A,523del,524del	L1c3a1b
Na13	16126C,16223T,16270T,16298C,16325C,16327T,71A,73G,249del,263G,290del,291del,309.1C,315.1C,489C,493G,523del,524del	C1b
Na14	16173T,16182C,16183C,16189C,16217C,16223T,16357C,16519C,73G,263G,309.1C,309.2C,315.1C,499A	B4b
Na15	16172C,16189C,16223T,16320T,16519C,73G,150T,152C,195C,263G,315.1C	L3e2b+152
Na16	16291T,16519C,263G,309.1C,315.1C	H2a5a1
Na17	16170R,16298C,72C,263G,309.1C,309.2C,315.1C	HV0
Na18	16111A,16223T,16234T,16249C,16278T,16294T,16295T,16390A,73G,143A,146C,152C,195C,263G,315.1C	L2a1+143+@16309
Na20	16176T,16218T,16519C,200G,251A,263G,309.1C,315.1C	H3an
Na21	16223T,16278T,16294T,16309G,16368C,16390A,16519C,73G,146C,152C,195C,263G,315.1C,524.1A,524.2C,	L2a1a1
Na22	16051G,16093C,16223T,16298C,16325C,16327T,16519C,73G,189G,194T,249del,263G,290del,291del,309.1C,315.1C,489C,523del,524del	C1d+194
Na23	16126C,16294T,16296T,16304C,16362C,16519C,73G,263G,309.1C,315.1C	T2b+16362
Na24	16172C,16189C,16223T,16320T,16519C,73G,150T,195C,263G,309.1C,315.1C	L3e2b
RE01	16178C,16183C,16189C,16193.1C,16193.2C,16217C,16519C,73G,263G,309.1C,309.2C,315.1C,499A	B4b
RE02	16051G,16189C,16234T,16270T,73G,146C,150T,263G,315.1C	U5b1b1+@16192
RE03	16223T,16239T,16325C,16362C,16519C,73G,210G,228A,263G,315.1C,489C,524.1A,524.2C	M21b+210
RE04	16168T,16182C,16183C,16189C,16217C,16249C,16312G,16344T,16519C,73G,152C,263G,271T,309.1C,315.1C,499A	B2b3a
RE05	16519C,263G,309.1C,315.1C	H2a2a
RE06	16189C,16298C,72C,263G,315.1C	HV0
RE07	16111T,16223T,16290T,16319A,16362C,64T,73G,146C,153G,235G,263G,309.1C,315.1C,485C,523del,524del	A2+(64)

Regarding the European (H, HV0, U5 and V), the Middle East (T2) and the North African (X1) lineages that could reflect putative Jewish ancestry, a search for similar Control Region haplotypes was performed in the EMPOP database (http://empop.org/). Matches were obtained for HV0 and T2 lineages with samples from Zamora (Spain) and Morocco, both places with a well-documented Jewish history.

Diversity measures estimated from the analysis of the mtDNA control region are in the same range of values calculated for other general Brazilian populations [156, 177, 178]: high levels of gene diversity were obtained (0.9985 ± 0.0067) for the entire Control Region, (excluding poli-C

regions); the mean number of pairwise differences was (16.022523 ± 7.309300) and the nucleotide diversity, averaged over loci, was (0.014179 ± 0.007189).

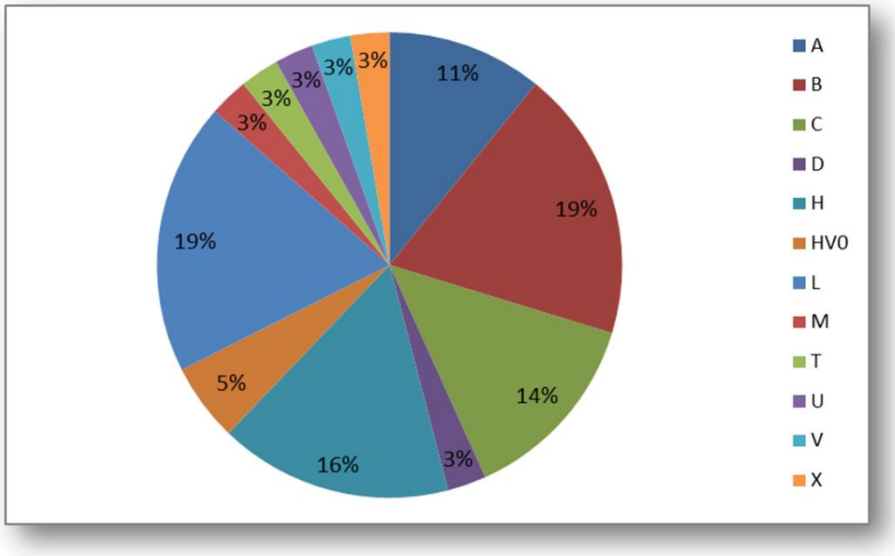


Figure 7. Mitochondrial DNA major haplogroup composition and corresponding relative frequencies for the total sample of the Brazilian Jewish communities.

In conclusion, estimates of parental contribution based on mitochondrial DNA and Y-chromosome markers in the Jewish population of Brazil can be summarized in the following **Table 3**:

	Parental contribution (%)					
	European	**African**		**Middle East**	**Amerindian**	**Others**
		Sub-Saharan	**North**			
mtDNA	27	18.9	2.7	2.7	45.9	2.7
Y Cr	45.8	0	12.5	37.5	0	4.2

Table 3. Estimates of parental contribution based on mtDNA and Y-chromosome (Y Cr) markers in the Brazilian Jews.

The sex-biased pattern observed for the present population, where a strong Sub-Saharan African and Amerindian female contribution contrasts with the European plus Middle Eastern and North African male input, reflects the disproportionate excess of male migrants, along with a strong introgression of Amerindian and African females into the Jewish communities. However,

this particular pattern has also been detected in other Brazilian groups other than Jews, e.g. [179], and is the outcome of the complex demographic movements of migration and colonization events that took place since the 16th century.

Enlarging the Jewish sample as well as further analysis using different genetic polymorphisms, such as autosomes and X chromosome linked markers and also fast evolving Y chromosome STRs, would surely contribute for a deeper knowledge of the Jewish population of Brazil and its relation to the Iberian communities.

CHAPTER V. DISCUSSION

This thesis focused on the characterization of the genetic pool of the Portuguese Crypto-Jewish descendants, representatives of the original Sephardic population that stayed in Portugal after the Expulsion Edit in 1496. Besides this, Sephardic Jewish communities of Brazil as well as presumed Jewish descendants from São Tomé islands were also surveyed.

Overall, the results obtained during this thesis allowed to reach the following main conclusions:

- The newly complete mtDNA genomes obtained, allowed the definition of five putative Sephardic founding lineages, namely HV0b, N1, T2b11, T2e and U2e. HV0b lineage shares a common private variant 8520A>G with Belmonte Jews, moreover, the NE Portuguese Jewish samples cluster together, sharing a more recent variant not previously described, 10644G>A, which seems to have arisen locally.

  Lineage N1 is shared between the Portuguese Jews and individuals from the neighbor regions, Zamora (Spain) and Miranda do Douro, both places with a well-known Jewish past besides an Ashkenazi Jew. Lineages carrying the 150C>T transition were not found in available databases and thus represent what could be a specific feature of the Sephardic Jews from NE Portugal.

  Variant 9181A>G within the T2b11 branch seems to be regionally specific, reflecting a Sephardic signature, given its absence from public databases. An additional motif, 4902A>G-8557G>A-16167C>T-16261C>T, was found exclusively among the Portuguese Jews.

  T2e has been described as a founding lineage in other Sephardic communities although it also includes Ashkenazim Jews and again samples from Zamora. The Portuguese Jews present two distinct variants inside T2e lineage, 13135G>A and 7133C>T, the latter not described until now.

  U2e1 lineages found in the Jewish population are shared with two individuals from Miranda do Douro and one Ashkenazi Jew.

- NE communities, despite preserving a distinctive lineage profile, displayed diversity levels similar to the host population, in sharp contrast to the strong founder effect found in the only Jewish Portuguese population reported before – the community of Belmonte.

- The estimates of both the diversity levels and the number of female effective-population founders, pointed at a stable size of the studied populations, in agreement with previous findings for the male counterparts.
- Signs of introgression from the host non-Jewish population were also detected.
- The simultaneous analysis of both autosomal and heterosomal markers revealed a sex biased demographic history, suggesting an asymmetry between female and male effective population sizes in the admixture process between Jews and non-Jews, not reported historically.
- High genetic diversity levels were also found both for autosomal as well as for the X-chromosome, despite a tendency for an allelic loss observed particularly with autosomal markers.
- Regarding the supposed Jewish descendants from São Tomé, despite the fact that almost 50% of the surnames of the participants in the study correspond to the most frequent surnames found in Inquisitional processes for Judaizing, a much stronger African ancestry component was found compared to the lower European fraction for all set of markers tested: the 46 Ancestry Informative Markers and 38 non-coding bi-allelic autosomal Indels showed 86% African and 14% European; while the lineage markers showed 80% African and 20% European for the Y chromosome and 95% African to 5% European (corresponding to a single female) for the mtDNA. Overall, no traces of Jewish ancestry were found in the European part of their genetic pool.
- The Jewish communities in Brazil showed high levels of genetic diversity for both the female and male lineages. A strong sex-biased pattern was detected showing a disproportionate excess of male migrants, along with a strong introgression of native and african females into the Jewish communities. Native lineages were absent in the male gene pool, which presented lineages from the Middle East (G e J2), Europe (R1a e R1b) and north Africa (E1b), that could be compatible with a Jewish ancestry. On the female counterpart, the vast majority of lineages found were of Native American and African origin. Nevertheless, traces of European lineages (H, HV0, U5 e V), North Africa (X1), and Middle East (T2) were also found. Further studies will be needed to clarify the demographic evolution of these communities. A greater sample size would be desirable along with the analysis of the recombining genomes (autosomal and X-linked markers).

As a whole, the results obtained during this thesis allowed a comprehensive reconstruction of the Sephardic Jewish history. Genetic data obtained, together with the historical available

information for each population was used in a multidisciplinary approach to clarify aspects of the complex demographic history of the Jewish Sephardic population. More Diaspora Iberian Jewish communities, particularly from North America and the North of Europe allied to a whole genome approach would surely contribute to complement what this work began to disclose.

CHAPTER VI. REFERENCES

1. Kaplan, Y., *Dicionário do Judaismo Português*. 2009, Editorial Presença: Lisboa. p. 174-176, 192-206.
2. Mea, E.d.A., *Dicionário do Judaismo Português*. 2009, Editorial Presença: Lisboa. p. 186-187.
3. Stuczinsky, C., *Dicionário do Judaismo Português*. 2009, Editorial Presença: Lisboa. p. 341-343.
4. Barnavi, É., *História Universal dos Judeus.* 1992, Lisboa.: Contexto.
5. Steinhardt, I., *Raízes dos Judeus em Portugal*. 1º ed. 2012, Lisboa: Nova Vega Lda. 179.
6. Levy-Coffman, E., *A mosaic of people: the Jewish story and a reassessment of the DNA evidence.* Journal of Genetic Genealogy, 2005. **1**: p. 12-33.
7. Barnes, I., *The Historical Atlas of Judaism*, ed. J. Bacon. 2009, London: Cartographica Press. 400.
8. Schwarz, S., *Os cristãos-novos em Portugal no século XX*. 1993: Instituto de Sociologia e Etnologia das Religiões, Universidade Nova de Lisboa.
9. Paulo, A., *Os judeus secretos em Portugal*. 1985: Editorial Labirinto.
10. Tavares, M.J.P.F., *A herança judaica em Portugal*. 2004, Lisboa: CTT Correios de Portugal.
11. Wilke, C.L. and J.F.C. da Costa, *História dos judeus portugueses*. 2009: Ediçoes 70.
12. Martins, J., *Portugal e os judeus: Judaísmo e anti-semitismo no século XX*. 2006, Lisboa: Vega.
13. Garcia, M.A., *Os judeus de Belmonte: os caminhos da memória*. 1993, Lisboa: Instituto de Sociologia e Etnologia das Religiões, Universidade Nova de Lisboa.
14. Pignatelli, M., *A comunidade israelita de Lisboa: o passado e o presente na construção da etnicidade dos judeus de Lisboa*. 2000: Universidade Técnica de Lisboa, Instituto Superior de Ciências Sociais e Políticas.
15. Azevedo, E.M. *Cristãos Novos No Nordeste Transmontano*. in *Los caminos de Cervantes y Sefarad: actas del II Congreso Internacional*. 1995. Zamora, Bragança, Puebla de Sanabria, Benavente: Asociación Caminos de Cervantes y Sefarad.
16. Carvalho, A.C., *Os judeus do desterro de Portugal*. 1999: Quetzal Editores.
17. Canelo, D.A., *The Last Crypto-Jews of Portugal*. 1990: Judaica Press.
18. Tavares, R., *A jóia do marranismo em Portugal - O misterioso Agnus Dei judaico-cristão de Carção*, in *Almocreve de Carção - Um retrato das Gentes de Carção*. 2015, Associação Cultural dos Almocreves de Carção: Bragança.
19. DellaPergola, S., *Jewish Diaspora.* International Encyclopedia of the Social & Behavioral Sciences, 2004: p. 7963-7969.
20. Mea, E., *O resgate dos meninos de S. Tomé em Oríon*, in *Literatura e História - Actas do Colóquio Internacional*. 2004: Porto. p. 25-39.
21. Lipiner, E., *Os baptizados em pé: Estudos acerca da origem e da luta dos Cristãos-Novos em Portugal*. 1998: Vega (Lisboa, PT).
22. de Albuquerque, L., *A ilha de São Tomé nos séculos XV e XVI*. 1989: Publicações Alfa.
23. Novinski, A., *Cristãos Novos na Bahia: 1624-1654*. São Paulo: Perspectiva.
24. Novinski, A., *Os cristãos-novos no Brasil colonial: reflexões sobre a questão do marranismo.* Tempo, 2001. **6**(11): p. 67-75.
25. Kaufman, T.N., *Novos personagens. Novas identidades. O marranismo contemporâneo em Pernambuco.* JUDAÍSMO E MODERNIDADE: SUAS MÚLTIPLAS INTER-RELAÇÕES, 2009: p. 134.
26. Futuyma, D.J., *Evolutionary Theory*, in *Human evolutionary biology*, M.P. Muehlenbein, Editor. 2010, Cambridge University Press. p. 3-16.

27. Relethford, J.H., *Human Population Genetics*, J.H. Relethford, Editor. 2012, John Wiley & Sons.
28. Crawford, M.H., *Foundations of Anthropological Genetics*, in *Anthropological genetics: theory, methods and applications*, M.H. Crawford, Editor. 2007, Cambridge University Press. p. 1-16.
29. Jobling, M.A., M. Hurles, and C. Tyler-Smith, *Human evolutionary genetics: origins, peoples & disease.* 2004.
30. Lynch, M., *The frailty of adaptive hypotheses for the origins of organismal complexity.* Proceedings of the National Academy of Sciences, 2007. **104**(suppl 1): p. 8597-8604.
31. Kimura, M., *Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles*.* Genetical research, 1968. **11**: p. 247-69.
32. Kimura, M., *Evolutionary rate at the molecular level.* Nature, 1968. **217**(5129): p. 624-626.
33. King, J.L. and T.H. Jukes, *Non-darwinian evolution.* Science, 1969. **164**(3881): p. 788-798.
34. Nei, M., Y. Suzuki, and M. Nozawa, *The neutral theory of molecular evolution in the genomic era.* Annual review of genomics and human genetics, 2010. **11**: p. 265-289.
35. Amorim, A., *Genetic markers: The interplay between concepts and technology in the Anthropological scene.* Recent Advances in Molecular Biology and Evolution: Applications to Biological Anthropology, 2007.
36. Corbett-Detig, R.B., D.L. Hartl, and T.B. Sackton, *Natural selection constrains neutral diversity across a wide range of species.* 2015.
37. Raff, R.J.M.a.J.A., *Genetics in Human Biology*, in *Human evolutionary biology*, M.P. Muehlenbein, Editor. 2010, Cambridge University Press. p. 48-73.
38. Slatkin, M., *Linkage disequilibrium—understanding the evolutionary past and mapping the medical future.* Nature Reviews Genetics, 2008. **9**(6): p. 477-485.
39. Meier, R.J., *The nature of human biological and genetic variability.* Physical (Biological) Anthropology Encyclopedia Of Life Support Systems (Eolss), 2010.
40. Holsinger, K.E. and B.S. Weir, *Genetics in geographically structured populations: defining, estimating and interpreting FST.* Nature Reviews Genetics, 2009. **10**(9): p. 639-650.
41. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.
42. Steiper, M.E., *DNA Markers of Human Variation*, in *Human evolutionary biology*, M.P. Muehlenbein, Editor. 2010, Cambridge University Press. p. 238-264.
43. Rohina Rubicz, P.M., and Michael Crawford, *Molecular Markers in Anthropological Genetic Studies*, in *Anthropological genetics: theory, methods and applications*, M. H.Crawford, Editor. 2007, Cambridge University Press. p. 141-186.
44. Consortium, I.H.G.S., *Finishing the euchromatic sequence of the human genome.* Nature, 2004. **431**(7011): p. 931-945.
45. Venter, J.C., et al., *The sequence of the human genome.* science, 2001. **291**(5507): p. 1304-1351.
46. Reich, D.E., et al., *Human genome sequence variation and the influence of gene history, mutation and recombination.* Nat Genet, 2002. **32**(1): p. 135-42.
47. Lander, E.S., *Initial impact of the sequencing of the human genome.* Nature, 2011. **470**(7333): p. 187-197.
48. Garrigan, D. and M.F. Hammer, *Reconstructing human origins in the genomic era.* Nature Reviews Genetics, 2006. **7**(9): p. 669-680.
49. Kidd, K.K., et al., *Understanding human DNA sequence variation.* Journal of Heredity, 2004. **95**(5): p. 406-420.

50. Nachman, M.W. and S.L. Crowell, *Estimate of the mutation rate per nucleotide in humans.* Genetics, 2000. **156**(1): p. 297-304.
51. Griffiths, A.J., Susan R. Wessler, Sean B. Carroll, John Doebley, *introduction to genetic analysis*. 10th ed. 2012, New York: W. H. Freeman and Company.
52. Antonarakis, S.E., *Human genome sequence and variation*, in *Vogel and Motulsky's Human Genetics*. 2010, Springer. p. 31-53.
53. Underhill, P.A. and T. Kivisild, *Use of Y chromosome and mitochondrial DNA population structure in tracing human migrations.* Annu. Rev. Genet., 2007. **41**: p. 539-564.
54. Hughes, J.F. and S. Rozen, *Genomics and genetics of human and primate y chromosomes.* Annual review of genomics and human genetics, 2012. **13**: p. 83-108.
55. Hallast, P., et al., *The Y-chromosome tree bursts into leaf: 13,000 high-confidence SNPs covering the majority of known clades.* Molecular biology and evolution, 2015. **32**(3): p. 661-673.
56. Pugach, I. and M. Stoneking, *Genome-wide insights into the genetic history of human populations.* Investigative genetics, 2015. **6**(1): p. 6.
57. Fregel, R., et al., *Carriers of Mitochondrial DNA Macrohaplogroup N Lineages Reached Australia around 50,000 Years Ago following a Northern Asian Route.* PloS one, 2015. **10**(6): p. e0129839.
58. Liu, N., et al., *Comparison of single-nucleotide polymorphisms and microsatellites in inference of population structure.* Bmc Genetics, 2005. **6**(Suppl 1): p. S26.
59. Haasl, R.J. and B.A. Payseur, *Multi-locus inference of population structure: a comparison between single nucleotide polymorphisms and microsatellites.* Heredity, 2011. **106**(1): p. 158-171.
60. Henn, B.M., et al., *Fine-scale population structure and the era of next-generation sequencing.* Human molecular genetics, 2010. **19**(R2): p. R221-R226.
61. Kayser, M. and P. de Knijff, *Improving human forensics through advances in genetics, genomics and molecular biology.* Nature Reviews Genetics, 2011. **12**(3): p. 179-192.
62. Li, H., G. Glusman, and H. Hu, *Relationship estimation from whole-genome sequence data.* PLoS genetics, 2014. **10**(1).
63. Putman, A.I. and I. Carbone, *Challenges in analysis and interpretation of microsatellite data for population genetic studies.* Ecology and evolution, 2014. **4**(22): p. 4399-4428.
64. Oliveira, E.J., et al., *Origin, evolution and genome distribution of microsatellites.* Genetics and Molecular Biology, 2006. **29**(2): p. 294-307.
65. Ellegren, H., *Microsatellites: simple sequences with complex evolution.* Nature reviews genetics, 2004. **5**(6): p. 435-445.
66. Sun, J.X., et al., *Microsatellites are molecular clocks that support accurate inferences about history.* Molecular biology and evolution, 2009. **26**(5): p. 1017-1027.
67. Schlötterer, C., *Evolutionary dynamics of microsatellite DNA.* Chromosoma, 2000. **109**(6): p. 365-371.
68. Guichoux, E., et al., *Current trends in microsatellite genotyping.* Molecular Ecology Resources, 2011. **11**(4): p. 591-611.
69. Weber, J.L., et al., *Human diallelic insertion/deletion polymorphisms.* The American Journal of Human Genetics, 2002. **71**(4): p. 854-862.
70. Mullaney, J.M., et al., *Small insertions and deletions (INDELs) in human genomes.* Human molecular genetics, 2010. **19**(R2): p. R131-R136.
71. Mills, R.E., et al., *An initial map of insertion and deletion (INDEL) variation in the human genome.* Genome research, 2006. **16**(9): p. 1182-1190.
72. Montgomery, S.B., et al., *The origin, evolution, and functional impact of short insertion–deletion variants identified in 179 human genomes.* Genome research, 2013. **23**(5): p. 749-761.

73. Narzisi, G. and M.C. Schatz, *The challenge of small-scale repeats for indel discovery.* Frontiers in bioengineering and biotechnology, 2015. **3**.
74. Mullikin, J., et al., *An SNP map of human chromosome 22.* Nature, 2000. **407**(6803): p. 516-520.
75. Pereira, R., et al., *A new multiplex for human identification using insertion/deletion polymorphisms.* Electrophoresis, 2009. **30**(21): p. 3682-3690.
76. Pereira, R., et al., *Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing.* PLoS One, 2012. **7**(1): p. e29684.
77. Phillips, C., *Forensic genetic analysis of bio-geographical ancestry.* Forensic Science International: Genetics, 2015.
78. Schaffner, S.F., *The X chromosome in population genetics.* Nature Reviews Genetics, 2004. **5**(1): p. 43-51.
79. Jobling, M.A., *The impact of recent events on human genetic diversity.* Philos Trans R Soc Lond B Biol Sci, 2012. **367**(1590): p. 793-9.
80. Schiffels, S. and R. Durbin, *Inferring human population size and separation history from multiple genome sequences.* Nature genetics, 2014.
81. Pritchard, J.K., M. Stephens, and P. Donnelly, *Inference of population structure using multilocus genotype data.* Genetics, 2000. **155**(2): p. 945-959.
82. Hubisz, M.J., et al., *Inferring weak population structure with the assistance of sample group information.* Molecular ecology resources, 2009. **9**(5): p. 1322-1332.
83. Behar, D.M., et al., *The genome-wide structure of the Jewish people.* Nature, 2010.
84. Kopelman, N.M., et al., *Genomic microsatellites identify shared Jewish ancestry intermediate between Middle Eastern and European populations.* BMC Genet, 2009. **10**: p. 80.
85. Listman, J.B., et al., *Identification of population substructure among Jews using STR markers and dependence on reference populations included.* BMC Genet, 2010. **11**: p. 48.
86. Campbell, C.L., et al., *North African Jewish and non-Jewish populations form distinctive, orthogonal clusters.* Proc Natl Acad Sci U S A, 2012. **109**(34): p. 13865-70.
87. Atzmon, G., et al., *Abraham's children in the genome era: major Jewish diaspora populations comprise distinct genetic clusters with shared Middle Eastern Ancestry.* Am J Hum Genet, 2010. **86**(6): p. 850-9.
88. Picornell, A., et al. *STR data for 13 loci from Jewish populations*. in *International Congress Series*. 2003. Elsevier.
89. Ellegren, H., *Sex-chromosome evolution: recent progress and the influence of male and female heterogamety.* Nature reviews genetics, 2011. **12**(3): p. 157-166.
90. Huynh, K.D. and J.T. Lee, *X-chromosome inactivation: a hypothesis linking ontogeny and phylogeny.* Nature Reviews Genetics, 2005. **6**(5): p. 410-418.
91. Arbiza, L., et al., *Contrasting X-linked and autosomal diversity across 14 human populations.* The American Journal of Human Genetics, 2014. **94**(6): p. 827-844.
92. Hammer, M.F., et al., *Sex-biased evolutionary forces shape genomic patterns of human diversity.* PLoS genetics, 2008. **4**(9): p. e1000202.
93. Emery, L.S., J. Felsenstein, and J.M. Akey, *Estimators of the human effective sex ratio detect sex biases on different timescales.* The American Journal of Human Genetics, 2010. **87**(6): p. 848-856.
94. Heyer, E., et al., *Sex-specific demographic behaviours that shape human genomic variation.* Molecular ecology, 2012. **21**(3): p. 597-612.
95. Diegoli, T.M., *Forensic typing of short tandem repeat markers on the X and Y chromosomes.* Forensic Science International: Genetics, 2015.

96. Gomes, C., et al., *Comparative evaluation of alternative batteries of genetic markers to complement autosomal STRs in kinship investigations: autosomal indels vs. X-chromosome STRs.* International journal of legal medicine, 2012. **126**(6): p. 917-921.
97. Pereira, V., et al., *Refining the genetic portrait of Portuguese Roma through X-chromosomal markers.* American journal of physical anthropology, 2012. **148**(3): p. 389-394.
98. Valente, C., et al., *Association between STRs from the X chromosome in a sample of Portuguese Gypsies.* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 391-393.
99. Pinto, J., et al., *Mirandese language and genetic differentiation in Iberia: a study using X chromosome markers.* Annals of human biology, 2014. **42**(1): p. 20-25.
100. Bachtrog, D., *Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration.* Nature Reviews Genetics, 2013. **14**(2): p. 113-124.
101. Lahn, B.T., N.M. Pearson, and K. Jegalian, *The human Y chromosome, in the light of evolution.* Nature Reviews Genetics, 2001. **2**(3): p. 207-216.
102. Skaletsky, H., et al., *The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes.* Nature, 2003. **423**(6942): p. 825-37.
103. Sun, S. and J. Heitman, *Should Y stay or should Y go: The evolution of non-recombining sex chromosomes.* BioEssays, 2012. **34**(11): p. 938-942.
104. Hughes, J.F., et al., *Strict evolutionary conservation followed rapid gene loss on human and rhesus Y chromosomes.* Nature, 2012. **483**(7387): p. 82-86.
105. Hughes, J.F., et al., *Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content.* Nature, 2010. **463**(7280): p. 536-539.
106. Rozen, S., et al., *Remarkably little variation in proteins encoded by the Y chromosome's single-copy genes, implying effective purifying selection.* The American Journal of Human Genetics, 2009. **85**(6): p. 923-928.
107. Jobling, M.A. and C. Tyler-Smith, *The human Y chromosome: an evolutionary marker comes of age.* Nat Rev Genet, 2003. **4**(8): p. 598-612.
108. Sayres, M.A.W., K.E. Lohmueller, and R. Nielsen, *Natural selection reduced diversity on human y chromosomes.* 2014.
109. de Knijff, P., *Messages through bottlenecks: on the combined use of slow and fast evolving polymorphic markers on the human Y chromosome.* The American Journal of Human Genetics, 2000. **67**(5): p. 1055-1061.
110. Underhill, P.A., et al., *The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations.* Ann Hum Genet, 2001. **65**(Pt 1): p. 43-62.
111. Karafet, T.M., et al., *New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree.* Genome Res, 2008. **18**(5): p. 830-8.
112. Oven, M., et al., *Seeing the wood for the trees: a minimal reference phylogeny for the human Y chromosome.* Human mutation, 2014. **35**(2): p. 187-191.
113. Larmuseau, M.H., et al., *Towards a consensus Y-chromosomal phylogeny and Y-SNP set in forensics in the next-generation sequencing era.* Forensic Science International: Genetics, 2015. **15**: p. 39-42.
114. Hammer, M.F., et al., *Y chromosomes of Jewish priests.* Nature, 1997. **385**(3).
115. Thomas, M.G., et al., *Origins of old testament priests.* Nature, 1998. **394**(6689): p. 138-140.
116. Tofanelli, S., et al., *Mitochondrial and Y chromosome haplotype motifs as diagnostic markers of Jewish ancestry: a reconsideration.* Frontiers in genetics, 2014. **5**.
117. Oefner, P.J., et al., *Genetics and the history of the Samaritans: Y-chromosomal microsatellites and genetic affinity between Samaritans and Cohanim.* Human biology, 2013. **85**(6): p. 825-857.

118. Behar, D.M., et al., *Contrasting patterns of Y chromosome variation in Ashkenazi Jewish and host non-Jewish European populations.* Hum Genet, 2004. **114**(4): p. 354-65.
119. Behar, D.M., et al., *Multiple origins of Ashkenazi Levites: Y chromosome evidence for both Near Eastern and European ancestries.* The American Journal of Human Genetics, 2003. **73**(4): p. 768-779.
120. Hammer, M.F., et al., *Jewish and Middle Eastern non-Jewish populations share a common pool of Y-chromosome biallelic haplotypes.* Proc Natl Acad Sci U S A, 2000. **97**(12): p. 6769-74.
121. Nebel, A., et al., *The Y chromosome pool of Jews as part of the genetic landscape of the Middle East.* Am J Hum Genet, 2001. **69**(5): p. 1095-112.
122. Shen, P., et al., *Reconstruction of patrilineages and matrilineages of Samaritans and other Israeli populations from Y-chromosome and mitochondrial DNA sequence variation.* Hum Mutat, 2004. **24**(3): p. 248-60.
123. Carmi, S., et al., *Sequencing an Ashkenazi reference panel supports population-targeted personal genomics and illuminates Jewish and European origins.* Nat Commun, 2014. **5**: p. 4835.
124. Ostrer, H., *A genetic profile of contemporary Jewish populations.* Nat Rev Genet, 2001. **2**(11): p. 891-8.
125. Ostrer, H. and K. Skorecki, *The population genetics of the Jewish people.* Hum Genet, 2013. **132**(2): p. 119-27.
126. Chinnery, P.F. and G. Hudson, *Mitochondrial genetics.* British medical bulletin, 2013. **106**(1): p. 135-159.
127. Chinnery, P.F., *Mitochondrial DNA in Homo sapiens*, in *Human Mitochondrial DNA and the Evolution of Homo sapiens*. 2006, Springer. p. 3-15.
128. Chinnery, P.F., *The transmission and segregation of mitochondrial DNA in Homo sapiens*, in *Human Mitochondrial DNA and the Evolution of Homo sapiens*. 2006, Springer. p. 17-29.
129. Schwartz, M. and J. Vissing, *Paternal inheritance of mitochondrial DNA.* New England Journal of Medicine, 2002. **347**(8): p. 576-580.
130. Stewart, J.B. and N.-G. Larsson, *Keeping mtDNA in shape between generations.* 2014.
131. Pakendorf, B. and M. Stoneking, *Mitochondrial DNA and human evolution.* Annu Rev Genomics Hum Genet, 2005. **6**: p. 165-83.
132. Ramos, A., et al., *Frequency and Pattern of Heteroplasmy in the Complete Human Mitochondrial Genome.* PLoS One, 2013. **8**(10): p. e74636.
133. Sobenin, I.A., et al., *Quantitative assessment of heteroplasmy of mitochondrial genome: perspectives in diagnostics and methodological pitfalls.* BioMed research international, 2014. **2014**.
134. Kivisild, T., *Maternal ancestry and population history from whole mitochondrial genomes.* Investigative genetics, 2015. **6**(1): p. 3.
135. Bandelt, H.-J., et al., *Estimation of mutation rates and coalescence times: some caveats*, in *Human Mitochondrial DNA and the Evolution of Homo sapiens*. 2006, Springer. p. 47-90.
136. Andrews, R.M., et al., *Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA.* Nat Genet, 1999. **23**(2): p. 147.
137. Anderson, S., et al., *Sequence and organization of the human mitochondrial genome.* Nature, 1981. **290**(5806): p. 457-65.
138. van Oven, M. and M. Kayser, *Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation.* Hum Mutat, 2009. **30**(2): p. E386-94.
139. Behar, D.M., et al., *A "copernican" reassessment of the human mitochondrial DNA tree from its root.* Am J Hum Genet, 2012. **90**(4): p. 675-84.

140. Tikochinski, Y., et al., *mtDNA polymorphism in two communities of Jews.* American journal of human genetics, 1991. **48**(1): p. 129.
141. Thomas, M.G., et al., *Founding mothers of Jewish communities: geographically separated Jewish groups were independently founded by very few female ancestors.* Am J Hum Genet, 2002. **70**(6): p. 1411-20.
142. Picornell, A., et al., *Mitochondrial DNA sequence variation in Jewish populations.* Int J Legal Med, 2006. **120**(5): p. 271-81.
143. Picornell, A., et al., *Mitochondrial DNA HVRI variation in Balearic populations.* Am J Phys Anthropol, 2005. **128**(1): p. 119-30.
144. Behar, D.M., et al., *Counting the founders: the matrilineal genetic ancestry of the Jewish Diaspora.* PLoS ONE, 2008. **3**(4): p. e2062.
145. Behar, D.M., et al., *MtDNA evidence for a genetic bottleneck in the early history of the Ashkenazi Jewish population.* Eur J Hum Genet, 2004. **12**(5): p. 355-64.
146. Behar, D.M., et al., *The matrilineal ancestry of Ashkenazi Jewry: portrait of a recent founder event.* Am J Hum Genet, 2006. **78**(3): p. 487-97.
147. Velez, C., et al., *The impact of Converso Jews on the genomes of modern Latin Americans.* Hum Genet, 2012. **131**(2): p. 251-63.
148. Bedford, F.L., *Sephardic signature in haplogroup T mitochondrial DNA.* Eur J Hum Genet, 2012. **20**(4): p. 441-8.
149. Bedford, F.L., et al., *Clarifying Mitochondrial DNA Subclades of T2e from Mideast to Mexico.* Journal of Phylogenetics & Evolutionary Biology, 2013. **1**(121): p. 121.
150. Cavalli-Sforza, L.L., P. Menozzi, and A. Piazza, *The history and geography of human genes*. 1994: Princeton university press.
151. Serebrenick, S., *Breve história dos judeus no Brasil*. 1962: Edições Biblos.
152. Carvalho, B.M., et al., *Mitochondrial DNA mapping of social-biological interactions in Brazilian Amazonian African-descendant populations.* Genetics and Molecular Biology, 2008. **31**(1): p. 12-22.
153. Cardena, M.M., et al., *Assessment of the relationship between self-declared ethnicity, mitochondrial haplogroups and genomic ancestry in Brazilian individuals.* 2013.
154. Palencia, L., et al., *Mitochondrial DNA diversity in a population from Santa Catarina (Brazil): predominance of the European input.* International journal of legal medicine, 2010. **124**(4): p. 331-336.
155. Paneto, G.G., et al., *A single multiplex PCR and SNaPshot minisequencing reaction of 42 SNPs to classify admixture populations into mitochondrial DNA haplogroups.* Mitochondrion, 2011. **11**(2): p. 296-302.
156. Sanches, N.M., et al., *Mitochondrial DNA control region diversity in a population from Espirito Santo state, Brazil.* Molecular biology reports, 2014. **41**(10): p. 6645-6648.
157. Silva, W.A., et al., *MtDNA haplogroup analysis of black Brazilian and sub-Saharan populations: implications for the Atlantic slave trade.* Human biology, 2006: p. 29-41.
158. Gonçalves, V.F., et al., *The phylogeography of African Brazilians.* Human heredity, 2008. **65**(1): p. 23-32.
159. Abe-Sandes, K., W.A. Silva Jr, and M.A. Zago, *Heterogeneity of the Y chromosome in Afro-Brazilian populations.* Human Biology, 2004: p. 77-86.
160. Assis, A., et al., *Hierarchical analysis of 15 Y-chromosome SNPs and demographic history of Afro-derived isolated communities in Alagoas, Brazil.* Forensic Science International: Genetics Supplement Series, 2011. **3**(1): p. e172-e173.
161. de Azevedo, D.A., et al., *Analysis of Y chromosome SNPs in Alagoas, Northeastern Brazil.* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 421-422.
162. Carvalho-Silva, D.R., et al., *The phylogeography of Brazilian Y-chromosome lineages.* The American Journal of Human Genetics, 2001. **68**(1): p. 281-286.

163. Carvalho-Silva, D.R., et al., *Y chromosome diversity in Brazilians: switching perspectives from slow to fast evolving markers.* Genetica, 2006. **126**(1-2): p. 251-260.
164. Oliveira, A.M., et al., *Analysis of Y chromosome lineages in a sample from Sub-Saharan Africa descendents in Rio de Janeiro.* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 442-443.
165. Silva, D.A., et al., *Y-chromosome genetic variation in Rio De Janeiro population.* American Journal of Human Biology, 2006. **18**(6): p. 829-837.
166. Gonçalves, F.T., et al., *The E180splice mutation in the GHR gene causing Laron syndrome: Witness of a Sephardic Jewish exodus from the Iberian Peninsula to the New World?* American Journal of Medical Genetics Part A, 2014. **164**(5): p. 1204-1208.
167. Brion, M., et al., *Hierarchical analysis of 30 Y-chromosome SNPs in European populations.* Int J Legal Med, 2005. **119**(1): p. 10-5.
168. Gomes, V., et al., *Digging deeper into East African human Y chromosome lineages.* Human genetics, 2010. **127**(5): p. 603-613.
169. Gusmao, A., et al., *A Perspective on the History of the Iberian Gypsies Provided by Phylogeographic Analysis of Y-Chromosome Lineages.* Annals of human genetics, 2008. **72**(2): p. 215-227.
170. Hammer, M.F. and S. Horai, *Y chromosomal DNA variation and the peopling of Japan.* Am J Hum Genet, 1995. **56**(4): p. 951-62.
171. Ramos, A., et al., *Human mitochondrial DNA complete amplification and sequencing: a new validated primer set that prevents nuclear DNA sequences of mitochondrial origin co-amplification.* Electrophoresis, 2009. **30**(9): p. 1587-93.
172. Excoffier, L. and H.E. Lischer, *Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows.* Mol Ecol Resour, 2010. **10**(3): p. 564-7.
173. Kloss-Brandstätter, A., et al., *HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups.* Human mutation, 2011. **32**(1): p. 25-32.
174. Röck, A.W., et al., *Concept for estimating mitochondrial DNA haplogroups using a maximum likelihood approach (EMMA).* Forensic Science International: Genetics, 2013. **7**(6): p. 601-609.
175. Sao-Bento, M., et al., *Y-chromosome SNP analysis in the Brazilian population of Sao Paulo state (Ribeirao Preto).* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 427-428.
176. Nogueiro, I., et al., *Portuguese Crypto-Jews: The Genetic Heritage of a Complex History.* Frontiers in Genetics, 2015. **6**.
177. Fridman, C., et al., *Haplotype diversity in mitochondrial DNA hypervariable region in a population of southeastern Brazil.* International journal of legal medicine, 2014. **128**(4): p. 589-593.
178. Bernardo, S., et al., *mtDNA ancestry of Rio de Janeiro population, Brazil.* Molecular biology reports, 2014. **41**(4): p. 1945-1950.
179. Guerreiro-Junior, V., et al., *Genetic signatures of parental contribution in black and white populations in Brazil.* Genetics and molecular biology, 2009. **32**(1): p. 1-11.

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO